

الحمد لله الذي
خلقنا من
الحمم



دانشگاه صنعتی شاهرود

دانشکده مهندسی عمران

پایان نامه کارشناسی ارشد راه و ترابری

تخمین ماتریس سفر با استفاده از هوش مصنوعی

نگارنده: ابوالفضل سجادی

استاد راهنما

دکتر قاسم زاده طهرانی

مهر ۱۴۰۰

شماره: ۵۵۰ / ت-ع

تاریخ: ۱۳۸۸ / ۸ / ۱۴

ویرایش:

باسمه تعالی

فرم های ارزشیابی پایان نامه کارشناسی ارشد

مربوط به ورودی های ۹۴ به بعد



دانشگاه گیلان

مدیریت تحصیلات تکمیلی

فرم شماره (۳) صورت جلسه نهایی دفاع از پایان نامه دوره کارشناسی ارشد

با نام و یاد خداوند متعال، ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد خانم **ابوالفضل سجادی** با شماره دانشجویی **۹۷۰۸۹۶۴** رشته **مهندسی عمران گرایش راه و ترابری** تحت عنوان **تخمین ماتریس سفر با استفاده از هوش مصنوعی** که در تاریخ **۱۴۰۰/۰۷/۲۶** با حضور هیئت محترم داوران در دانشگاه صنعتی شاهرود برگزار شد به شرح ذیل اعلام می گردد:

☒ الف) درجه عالی: نمره ۲۰ - ۱۹	☐ ب) درجه خیلی خوب: نمره ۱۸/۹۹ - ۱۸
☐ ج) درجه خوب: نمره ۱۷/۹۹ - ۱۶	☐ د) درجه متوسط: نمره ۱۵/۹۹ - ۱۴
☐ ه) کمتر از ۱۴ غیر قابل قبول و نیاز به دفاع مجدد دارد	
نوع تحقیق: ☐ نظری ☐ عملی	

عضو هیئت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء
۱- استاد راهنمای اول	دکتر حسین قاسمزاده طهرانی	استادیار	
۲- استاد راهنمای دوم			
۳- استاد مشاور			
۴- استاد داور اول	دکتر ایمان آقایان	دانشیار	
۵- استاد داور دوم	دکتر عبدالاحد جویانی	استادیار	
۶- نماینده تحصیلات تکمیلی	دکتر عماد محجوبی	استادیار	

نام و نام خانوادگی رئیس دانشکده:

تاریخ و امضاء مهر دانشکده:

تقدیم اثر

تقدیم به مهربان فرشتگانی که:

لحظات ناب باور بودن، لذت و غرور دانستن، جسارت خواستن، عظمت رسیدن و تمام تجربه‌های

یکتا و زیبای زندگی‌ام، مدیون حضور سبز آن‌هاست

تقدیم به خانواده‌ی عزیزم.

تشکر و قدردانی

از خدای متعال سپاس گزارم که توفیق کسب علم و دانش را به من عطا فرمود تا بتوانم این مرحله از علم‌آموزی را با موفقیت به پایان برسانم. از خانواده‌ی عزیزم به خاطر حمایت‌ها و محبت‌های بی‌دریغی که نسبت به من داشته و دارند، کمال تشکر و سپاس را دارم. از استاد ارجمندم، جناب آقای **دکتر قاسم زاده طهرانی** به خاطر راهنمایی‌های ارزشمند و زحمات ایشان در کلیه‌ی مراحل انجام پایان‌نامه تقدیر و تشکر می‌نمایم.

در انتها از تمامی اساتید محترم دانشکده‌ی عمران دانشگاه صنعتی شاهرود که توفیق شاگردی‌شان را داشتم، سپاس‌گزاری نموده از خدای متعال سلامت و توفیق برای ایشان مسألت دارم.

ابوالفضل سجادی

مهر ۱۴۰۰

تعهد نامه

این جانب **ابوالفضل سجادی** دانشجوی دوره‌ی کارشناسی ارشد رشته‌ی مهندسی عمران گرایش راه و ترابری دانشکده‌ی مهندسی عمران دانشگاه صنعتی شاهرود نویسنده‌ی پایان‌نامه‌ی **تخمین ماتریس سفر با استفاده از هوش مصنوعی تحت راهنمایی دکتر قاسم زاده طهرانی** متعهد می‌شوم.

- تحقیقات در این پایان‌نامه توسط این جانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های محققان دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در پایان‌نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می‌باشد و مقالات مستخرج بانام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان‌نامه تأثیرگذار بوده‌اند در مقالات مستخرج از پایان‌نامه رعایت می‌گردد.
- در کلیه مراحل انجام این پایان‌نامه، در مواردی که از موجود زنده (یا بافت‌های آن‌ها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان‌نامه، در مواردی که به حوزه‌ی اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری، ضوابط و اصول اخلاق انسانی رعایت شده است.

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته شده) متعلق به دانشگاه صنعتی شاهرود می‌باشد. این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در پایان‌نامه بدون ذکر مرجع مجاز نمی‌باشد.

چکیده

از مهمترین پارامترها در مطالعات و مدیریت حمل و نقل، تعداد سفر بین مبدا و مقصد است که با استفاده از ماتریس سفر نشان داده می‌شود. از آنجا که تهیه ماتریس سفر نیازمند صرف زمان و هزینه بسیار زیادی است در سال‌های اخیر، روش‌های متعددی جهت تخمین ماتریس سفر ابداع و استفاده شده است. در این تحقیق هدف آن است تا با استفاده از هوش مصنوعی بتوان با استفاده از جمعیت ساکن در نواحی مبدا و مقصد و سطح زیربنای کاربری اراضی واقع در هر ناحیه، تخمین مناسبی از ماتریس سفر ارائه نمود. بدین منظور ابتدا تحقیقات پیشین بررسی شدند و نتیجه حاصل شد که می‌توان با استفاده از هوش مصنوعی تحقیقی مشابه با موضوع این تحقیق را به سرانجام رسانید. پس از آن داده‌های مربوط به سرشماری جمعیت و همچنین سطح زیربنای کاربری اراضی موجود در نواحی شهرداری در شهرهای مشهد و اصفهان جمع آوری شدند و پس از آماده سازی داده های خام، داده های آماده شده در دو حالت داده‌های نرمال سازی شده و داده های بدون نرمال سازی با استفاده از روش‌های (۱) شبکه عصبی جعبه ابزار نرم افزار MATLAB 2013 (۲) شبکه عصبی چندلایه بدون محدودیت داده‌های اعتبارسنجی (۳) شبکه عصبی با حذف داده‌های نویزی (۴) شبکه عصبی پیشخور جهت تخمین ماتریس سفر استفاده شدند.

در این تحقیق مناسبترین نتایج هنگام استفاده از داده های نرمال سازی شده در شبکه عصبی چندلایه بدون لحاظ کردن محدودیت اعتبارسنجی داده ها حاصل شد.

کلمات کلیدی: سفر، ماتریس سفر، تخمین، هوش مصنوعی، شبکه عصبی، MATLAB 2013، نرمال سازی داده‌ها.

فهرست مطالب

فهرست جداول.....	ل
فهرست اشکال.....	ن
فصل ۱ : مقدمه.....	۱
۱-۱ مقدمه.....	۲
۱-۲ ماتریس سفر.....	۳
۱-۳ روش های برآورد ماتریس سفر.....	۴
۱-۴ کاربرد هوش مصنوعی در تخمین ماتریس سفر.....	۵
۱-۵ ضرورت موضوع.....	۶
۱-۶ نوآوری موضوع.....	۶
۱-۷ فرضیات تحقیق.....	۶
۱-۸ فرضیه های تحقیق.....	۷
۱-۹ ساختار پایان نامه.....	۷
فصل ۲ : پیشینه تحقیق.....	۹
۲-۱ مقدمه.....	۱۰
۲-۲ معرفی موضوع.....	۱۰
۲-۳ روشهای تهیه ماتریس سفر.....	۱۱
۲-۳-۱ آمارگیری.....	۱۱
۲-۳-۲ مدلهای تخمینی.....	۱۲
۲-۴ مدل های توزیع سفر.....	۱۳
۲-۴-۱ مدل رشد:.....	۱۳
۲-۴-۲ مدل دیترویت:.....	۱۴
۲-۴-۳ مدل جاذبه:.....	۱۴
۲-۵ مدل تخمین ماتریس سفر با استفاده از حجم کمان.....	۱۵
۲-۵-۱ مدل کمترین مربعات.....	۱۷
۲-۵-۲ روش بیشترین آنتروپی (بیشترین جابجایی).....	۱۸

۲۰	 مدل بیشترین درستنمایی (maximum likely hood)
۲۱	 روش حداقل اطلاعات (minimum information)
۲۲	 مقاله Hichem Omrani
۲۴	 مقاله Gusri Yaldi
۳۰	 مقاله Albertengo
۳۷	 ۹- نتیجه گیری
۳۹	 فصل ۳: مبانی هوش مصنوعی
۴۰	 ۱-۳ مقدمه
۴۰	 ۲-۳ شبکه عصبی مصنوعی
۴۱	 ۱-۲-۳ مدل ریاضی نرون
۴۲	 ۲-۲-۳ توابع محرک
۴۵	 ۳-۲-۳ ساختار شبکه های عصبی
۴۷	 ۱-۳-۲-۳ شبکه های چند لایه
۴۸	 ۴-۲-۳ لایه خروجی
۴۹	 ۳-۳ روش های تخمین مبتنی بر هوش مصنوعی
۴۹	 ۱-۳-۳ مقدمه
۴۹	 ۲-۳-۳ شبکه عصبی
۵۰	 ۱-۲-۳-۳ یادگیری بدون نظارت
۵۰	 ۱-۲-۳-۳ یادگیری با نظارت
۵۱	 ۲-۲-۳-۳ یادگیری تقویتی
۵۱	 ۳-۳-۳ انواع شبکه عصبی
۵۲	 ۱-۳-۳-۳ شبکه عصبی پیش خور
۵۲	 ۲-۳-۳-۳ شبکه عصبی انتشار برگشتی
۵۲	 ۴-۳-۳ ویژگی های روش های شبکه عصبی
۵۳	 ۵-۳-۳ زمینه های کاربردی روش های شبکه عصبی
۵۳	 ۶-۳-۳ نرم افزارهای متداول برای محاسبات ANN
۵۵	 فصل ۴: روش تحقیق
۵۶	 ۱-۴ مقدمه

۲-۴	جمع آوری داده ها	۵۶
۱-۲-۴	جمع آوری داده‌های شهر مشهد	۵۶
۲-۲-۴	جمع آوری داده های شهر اصفهان	۵۷
۳-۴	آماده‌سازی داده‌ها	۶۰
۱-۳-۴	طبقه‌بندی کاربری اراضی	۶۱
۲-۳-۴	طبقه بندی سرشماری جمعیتی	۶۳
۳-۳-۴	آماده سازی داده های شهر مشهد	۶۴
۴-۳-۴	آماده سازی داده های شهر اصفهان	۶۶
۵-۳-۴	ضریب همبستگی میان داده ها	۶۹
۶-۳-۴	بررسی ضریب همبستگی پارامترها در نرم افزار SPSS	۷۱
۷-۳-۴	بررسی ضریب همبستگی در نرم افزار MATLAB	۷۳
۸-۳-۴	کد های استفاده شده	۷۴
۹-۳-۴	آماده سازی فایل اکسل ورودی	۸۰
فصل ۵:	مدلسازی و نتایج تحقیق	۸۳
۱-۵	مقدمه	۸۴
۲-۵	تحلیل نمودارهای خروجی	۸۴
۱-۲-۵	داده های شهر مشهد	۸۶
۱-۱-۲-۵	شبکه عصبی جعبه ابزار متلب	۸۷
۲-۱-۲-۵	شبکه عصبی چندلایه بدون داده های اعتبار سنجی	۸۹
۳-۱-۲-۵	شبکه عصبی با حذف داده های نویزی	۹۱
۲-۲-۵	داده های شهر اصفهان	۹۲
۱-۲-۲-۵	شبکه عصبی چندلایه بدون داده های اعتبار سنجی	۹۴
۲-۲-۲-۵	شبکه عصبی با حذف داده های نویزی	۹۶
۳-۲-۲-۵	شبکه عصبی پیشخور	۹۷
۳-۵	مقایسه ماتریس سفر تخمین‌زده شده و اولیه	۹۸
۱-۳-۵	ماتریس سفر تخمین‌زده شده شهر مشهد	۹۸
۲-۳-۵	ماتریس سفر تخمین‌زده شده شهر اصفهان	۱۰۱
فصل ۶:	نتایج	۱۰۵

۱-۶	مقدمه	۱۰۶
۲-۶	نتایج	۱۰۷
۳-۶	پیشنهادهای برای ادامه تحقیق:	۱۰۸
۱۰۹	مراجع	

فهرست جداول

جدول ۱-۲	تحقیقات انجام شده در زمینه مدل های تخمین جدول سفر	۱۶
جدول ۲-۲	مقایسه نتایج	۲۴
جدول ۳-۲	بلوک بندی داده ها	۲۵
جدول ۴-۲	عملکرد آزمایش شبکه عصبی آموزش داده شده با الگوریتم های متفاوت	۲۶
جدول ۵-۲	خلاصه ی خطای تخمین جریان وسایل نقلیه بوسیله یادگیری عمیق در داده های مختلف با پیش پردازش های متفاوت	۳۵
جدول ۱-۳	توابع محرک با علائم قراردادی	۴۲
جدول ۱-۴	بخشی از ماتریس سفر سال ۱۳۹۵ اصفهان بر اساس نواحی شهرداری	۵۸
جدول ۵-۴	بخشی از فایل سرشماری اصفهان در سال ۱۳۹۵ بر اساس ۱۵ ناحیه شهرداری	۵۸
جدول ۶-۴	کاربری اراضی موجود در نواحی مختلف شهرداری اصفهان بر اساس درصد	۵۹
جدول ۷-۴	سطح نواحی اصفهان و کاربری اراضی بر اساس هکتار	۶۰
جدول ۸-۴	ماتریس نهایی سفر مشهد بر اساس نواحی شهرداری	۶۵
جدول ۹-۴	داده های آماری سرشماری شهر مشهد سال ۱۳۹۵	۶۶
جدول ۱۰-۴	سطح زیربنای کاربری اراضی داخل هر ناحیه شهرداری شهر مشهد در سال ۱۳۹۵	۶۶
جدول ۱۱-۴	ماتریس سفر اصلی شهر اصفهان مربوط به سال ۱۳۹۵	۶۸
جدول ۱۲-۴	ماتریس سفر استفاده شده در این تحقیق، شهر اصفهان مربوط به سال ۱۳۹۵	۶۸
جدول ۱۳-۴	داده های سرشماری جمعیتی سال ۱۳۹۵ شهر اصفهان	۶۹
جدول ۱۴-۴	سطح زیربنای کاربری اراضی مختلف (بر اساس هکتار) در شهر اصفهان در سال ۱۳۹۵	۶۹
جدول ۱۵-۴	شیوه تفسیر شدت رابطه در همبستگی پیرسون	۷۱
جدول ۱۶-۴	دو سطر و پنج ستون از خروجی نرم افزار SPSS پارامترهای شهر اصفهان	۷۳
جدول ۱۷-۴	بخشی از جدول همبستگی های خروجی نرم افزار SPSS داده های شهر اصفهان	۷۳
جدول ۱۸-۴	بخشی از فایل ورودی به شبکه عصبی چندلایه بدون محدودیت داده اعتبارسنجی	۷۶
جدول ۱۹-۴	ماتریس سفر شهر اصفهان سال ۱۳۹۵	۸۰
جدول ۲۰-۴	بخشی از ماتریس ورودی به نرم افزار متلب جهت فرآیند تخمین	۸۱
جدول ۱-۵	ماتریس سفر Target شهر مشهد	۹۹

- جدول ۵-۲ ماتریس سفر OutPut شهر مشهد ۱۰۰
- جدول ۵-۳ ماتریس سفر Target شهر اصفهان ۱۰۱
- جدول ۵-۴ ماتریس سفر OutPut شهر اصفهان ۱۰۲

فهرست اشکال

- شکل ۱-۲ : معماری شبکه عصبی در این تحقیق ۲۵
- شکل ۲-۲ : واریانس نتایج ۲۷
- شکل ۳-۲ : تعداد تجمعی سفرهای برآورد شده ۲۷
- شکل ۴-۲ : رابطه رگرسیون تعداد سفرهای تخمین زده شده ۲۸
- شکل ۵-۲ : رابطه رگرسیون برآورد سفرهای جذب شده ۲۸
- شکل ۶-۲ : رابطه رگرسیون برآورد سفرهای تولید شده ۲۹
- شکل ۷-۲ : موقعیت سنسورهای برداشت اطلاعات ترافیکی در سطح شهر تورن ۳۰
- شکل ۸-۲ : جریان و سرعت ترافیک در شهر تورن بر اساس نوع روز ۳۱
- شکل ۹-۲ : فلوچارت پردازش داده ها بوسیله هوش مصنوعی ۳۲
- شکل ۱۰-۲ : مثالی از پیش پردازش پنجره های سری زمانی ۳۳
- شکل ۱۱-۲ : مثالی از جابجایی تصادفی جدول داده های پنجره ای ۳۴
- شکل ۱۲-۲ : شش حالت مختلف برای پیش پردازش داده ها در این مقاله ۳۵
- شکل ۱۳-۲ : جریان واقعی و جریان تخمین زده شده در یک روز هفته (a) با پیش پردازش میانگین متحرک و (b) بدون پیش پردازش میانگین متحرک ۳۶
- شکل ۱۴-۲ : جریان واقعی و جریان تخمین زده شده در یک روز آخر هفته (شنبه یکشنبه) با اعمال پیش پردازش میانگین متحرک بر داده ها ۳۷
- شکل ۱-۳ : ساختار یک نرون تک ورودی ۴۱
- شکل ۲-۳ : تابع محرک خطی ۴۳
- شکل ۳-۳ : تابع محرک آستانه ای دو مقدار حدی ۴۴
- شکل ۴-۳ : تابع محرک زیگموئید ۴۵
- شکل ۵-۳ : شبکه تک لایه با s نرون ۴۶
- شکل ۶-۳ : فرم فشرده یا ماتریسی شبکه تک لایه با s نرون ۴۶
- شکل ۷-۳ : شبکه پیش خور دو لایه ۴۷

- شکل ۳-۸ فرم ساده شده شبکه پیش‌خور سه لایه ۴۸
- شکل ۳-۹ : (الف) یادگیری بدون نظارت و (ب) یادگیری با نظارت ۵۱
- شکل ۴-۱ ناحیه بندی شهر مشهد بر اساس نواحی ۲۵۳ گانه ترافیکی ۶۴
- شکل ۴-۲ ناحیه بندی شهر مشهد بر اساس نواحی سیزده گانه ی شهرداری ۶۵
- شکل ۴-۳ ناحیه‌بندی شهرداری اصفهان ۶۷
- شکل ۵-۱ نتایج تخمین ماتریس سفر شهر مشهد با استفاده از شبکه عصبی پیش فرض جعبه ابزار متلب بوسیله داده های با ضریب همبستگی کمتر از ۰/۹ ۸۹
- شکل ۵-۲ نتایج تخمین ماتریس سفر شهر مشهد با استفاده از شبکه عصبی طراحی شده بدون محدودیت داده های اعتبارسنجی بوسیله داده های با ضریب همبستگی کمتر از ۰/۹ ۹۰
- شکل ۵-۳ شبکه عصبی چندلایه بدون داده های اعتبارسنجی (داده های بدون نرمال سازی) داده های با ضریب همبستگی ۰/۷ ۹۱
- شکل ۵-۴ نتایج تخمین ماتریس سفر شهر مشهد با استفاده از شبکه عصبی طراحی شده دارای قابلیت حذف داده‌های نویزی بوسیله داده های با ضریب همبستگی کمتر از ۰/۶ ۹۲
- شکل ۵-۵ نتایج تخمین ماتریس سفر شهر اصفهان با استفاده از شبکه عصبی طراحی شده بدون محدودیت داده های اعتبارسنجی بوسیله داده های با ضریب همبستگی کمتر از ۰/۸ ۹۵
- شکل ۵-۶ شبکه عصبی چندلایه بدون داده های اعتبارسنجی (داده های بدون نرمال سازی) داده های با ضریب همبستگی ۰,۹ ۹۵
- شکل ۵-۷ نتایج تخمین ماتریس سفر شهر اصفهان با استفاده از شبکه عصبی طراحی شده دارای قابلیت حذف داده‌های نویزی بوسیله داده های با ضریب همبستگی کمتر از ۰/۸ ۹۷
- شکل ۵-۸ نتایج تخمین ماتریس سفر شهر اصفهان با استفاده از شبکه عصبی پیش‌خور بوسیله داده های با ضریب همبستگی کمتر از ۰/۹ ۹۸

فصل ۱ : مقدمه

۱-۱ مقدمه

سفر عبارت است از حرکت از مکانی به نام مبدا (Origin) به مکانی به نام مقصد (Destination) که با هدف مشخص و وسیله خاص صورت می‌گیرد. آنچه در بحث تحلیل تقاضای حمل و نقل بیشتر مورد توجه قرار دارد، پیش‌بینی تعداد سفر است. در فرآیند پیش‌بینی تعداد سفر، هدف تحلیل‌گر آن است که برآورد کند در سال افق طرح در محدوده شهری مورد مطالعه، در هر ناحیه چه تعداد سفر ایجاد (تولید یا جذب) می‌شود، توزیع این سفرها چگونه است و از میان سفرهای انجام شده بین هر دو ناحیه سهم هر وسیله‌ی نقلیه در سرویس‌دهی این سفرها چقدر است و از هر یک از مسیرهای شبکه حمل و نقل چه حجمی عبور می‌کند (مدل چهار مرحله‌ای).

از ابزارهای اصلی در تحلیل تقاضای حمل و نقل ماتریس سفر می‌باشد. ماتریس سفر ماتریسی مربعی است و ابعاد آن به اندازه‌ی تعداد نواحی منطقه‌ی مورد مطالعه می‌باشد و نشان می‌دهد از هر ناحیه مانند i به هر ناحیه مانند j چه تعداد سفر صورت می‌گیرد. در صورت در اختیار داشتن ماتریس‌های مبدا-مقصد نزدیک به واقعیت، علاوه بر توصیف وضع موجود میتوان مدل‌های واقع بینانه‌تری با آنها ایجاد کرد. جهت ایجاد ماتریس سفر بصورت متداول از روش آمارگیری پرسشنامه‌ای سفرهای بین نواحی مبدا-مقصد استفاده می‌شود که بسیار پرهزینه است. اما می‌توان بصورت تقریبی نیز ماتریس سفر فعلی و ماتریس سفر سال افق را از ماتریس سفر مربوط به سال‌های پیشین تخمین زد که در هزینه‌های مطالعات حمل و نقل کاهش چشمگیری ایجاد می‌شود و در صورت دقت بالاتر در تخمین، ماتریس تخمین زده شده به واقعیت شباهت بیشتری خواهد داشت و در نتیجه می‌توان بر اساس آن با دقت بیشتری تصمیمات مربوط به مدیریت حمل و نقل در سال افق را اتخاذ کرد. لذا به روش‌هایی برای تخمین ارزان‌تر و با دقت بالاتر نیاز فراوان وجود دارد.

روش تهیه ماتریس سفر برای تهیه ماتریس مبدا-مقصد بصورت متداول از روش آمارگیری پرسشنامه‌ای و روش تخمین ماتریس سفر بوسیله شمارش ترافیک استفاده می‌شود. اخیرا با پیشرفت

تکنولوژی می‌توان بوسیله هوش مصنوعی به تخمین ماتریس سفر سال افق طرح، بر اساس ماتریس سفر سال‌های پیشین اقدام نمود.

۲-۱ ماتریس سفر

یک دستاورد مهم و با ارزش از آمارگیری مبداء-مقصد که با صرف وقت و هزینه فراوان، پس از جمع‌آوری اطلاعات سفر و تحلیل آنها بدست می‌آید، ماتریس مبداء-مقصد می‌باشد. این ماتریس نشان می‌دهد از هر ناحیه مانند i به هر ناحیه مانند j چه تعداد سفر صورت می‌گیرد. در معادله ۱-۱ ماتریس سفر بصورت ریاضی بیان شده است که در این معادله T ماتریس سفر و t_{ij} تعداد سفر بین هر دو ناحیه است.

$$T = [t_{ij}]_{m \times n} \quad \text{معادله ۱-۱}$$

در ماتریس سفر، جمع درایه‌های هر سطر مانند سطر i ام، تولید سفر ناحیه i را نشان می‌دهد که در این معادله P_i تعداد سفر تولید شده از ناحیه i است (معادله ۲-۱):

$$P_i = \sum_{j=1}^n t_{ij} \quad \text{معادله ۲-۱}$$

در ماتریس سفر، جمع درایه‌های هر ستون مانند ستون j ام نیز، تعداد سفرهای جذب شده به ناحیه j را نشان می‌دهد که در این معادله a_j تعداد سفر جذب شده به ناحیه j است (معادله ۳-۱):

$$a_j = \sum_{i=1}^n t_{ij} \quad \text{معادله ۳-۱}$$

در ماتریس سفر، جمع سفرهای تولید شده با تعداد سفرهای جذب شده، برابر و برابر با تعداد کل سفرهای انجام شده می‌باشد (معادله ۴-۱):

$$\sum_{i=1}^n P_i = \sum_{j=1}^n a_j = T \quad \text{معادله ۴-۱}$$

می‌توان بر اساس هدف سفر ماتریس سفرهای جداگانه‌ای تهیه کرد مانند ماتریس سفرهای تفریحی، سفرهای تحصیلی، سفرهای کاری و... . همچنین می‌توان بر اساس زمان سفر ماتریس سفرهای ساعت اوج و سفرهای غیر ساعت اوج را تهیه نمود.

۱-۳ روش‌های برآورد ماتریس سفر

پرهزینه‌ترین و زمانبرترین روش جهت ایجاد ماتریس سفر، آمارگیری جامع حمل و نقل می‌باشد. این آمارگیری مبدأ-مقصد با توجه به وقت‌گیر و پرهزینه بودن هر ۱۰ الی ۲۰ سال انجام می‌شود و در این روش اطلاعات جامعی در خصوص تعداد و نوع سفرهای میان نواحی مورد مطالعه بدست می‌آید که از آنها جهت تشکیل ماتریس مبدأ-مقصد استفاده می‌شود.

در صورتی که ماتریس سفر از گذشته موجود باشد می‌توان بر اساس اطلاعات موجود در آن و بوسیله روش‌های تخمین ماتریس سفر با توجه به ماتریس سفر سال‌های پیشین، ماتریس سفر کنونی و ماتریس سفر سال افق طرح را از روی اطلاعات موجود تخمین زد. برای تخمین ماتریس سفر در ابتدا به نمونه برداری از جامعه‌ی مورد مطالعه نیاز داریم. به طور کلی هدف از نمونه‌برداری کسب اطلاعاتی درباره ویژگی افراد جامعه بدون صرف هزینه‌های زیاد است. از این رو در نمونه‌برداری ابتدا می‌بایست نوع اطلاعاتی که قرار است ارزیابی شده و به جامعه نسبت داده شود، تعیین گردد. این اطلاعات با توجه به هدف کلی مطالعه و پارامترهای مورد نیاز برای مطالعه تعیین شده و اغلب به آنها پارامترهای جامعه گفته می‌شود. استنباط‌های آماری را می‌توان راجع به میانگین این پارامترها انجام داد. برای نمونه درآمد و نرخ سفر خانوار، پارامترهای جامعه به حساب می‌آیند. برای تعیین پارامتر مناسب در یک مطالعه به تجربه کافی و اطلاع از نوع کاربرد پارامتر در استنباط، تحلیل و یا مدل‌سازی نیاز است.

۴-۱ کاربرد هوش مصنوعی در تخمین ماتریس سفر

اصطلاح هوش مصنوعی برای اولین بار توسط جان ماکرتی (پدر علم ماشین‌های هوشمند) استفاده شد. هوش مصنوعی الگوریتم‌های محاسباتی متفاوتی از قبیل منطق فازی، شبکه عصبی مصنوعی، الگوریتم ژنتیک، یادگیری ماشین و الگوریتم‌های متفاوت دیگر را شامل می‌شود که انتخاب الگوریتم مناسب جهت انجام محاسبات بر اساس نوع داده‌های ورودی و خروجی و دقت مورد نیاز انجام می‌شود. هوش مصنوعی قابلیت‌های فراوانی دارد که از مهمترین کاربردهای مورد توجه آن در این تحقیق میتوان به قابلیت استنتاج و استدلال و تشخیص الگو و بازشناسی الگو برای پاسخگویی به مسائل بر اساس دانش قبلی اشاره کرد. از جمله مطالعات مرتبط با سیستم‌های حمل‌ونقل که تاکنون به کمک هوش مصنوعی انجام شده‌اند می‌توان به مطالعات در زمینه‌های بررسی رفتار راننده، وسایل نقلیه خودمختار، برآورد پارامترهای موثر بر شرایط مختلف ترافیکی، تشخیص و طبقه‌بندی خودروها بوسیله‌ی هوش مصنوعی در قالب پردازش تصویر، تحلیل الگوی ترافیک، پیش‌بینی ترافیک، تحلیل سیاست‌گذاری بر حمل‌ونقل و تاثیر آن بر اقتصاد، مدیریت و بهینه‌سازی سیستم‌های حمل و نقل هوایی، مدیریت و بهینه‌سازی حمل‌ونقل دریایی، مدیریت و بهینه‌سازی عملکرد سیستم مترو شهری و مدیریت و بهینه‌سازی سیستم‌های کنترل ترافیک اشاره کرد. بوسیله هوش مصنوعی شما می‌توانید از روش های گوناگون، ماتریس سفر را برآورد کنید و در هر روش شما می‌توانید از الگوریتم های ساده و پیچیده و متفاوت در حالت‌های تک لایه و چند لایه کمک بگیرید که بر اساس نوع داده‌ها (کیفی و کمی) و دقت مورد نیاز، الگوریتم مناسب طراحی و برنامه نویسی می‌شود.

در تخمین ماتریس سفر به این روش، پس از تعیین الگوریتم مناسب، با وارد کردن داده های اولیه، هوش مصنوعی شما، ارتباط پارامتر های ورودی را با تعداد سفر های میان نواحی مورد مطالعه، تشخیص داده و می‌توانید با اعمال شرایط سال افق طرح بر روی آن پارامترها، تعداد سفرهای بین نواحی را مشخص کنید.

۱-۵ ضرورت موضوع

از مهمترین زیرساخت‌های هر کشور می‌توان به شبکه حمل و نقل آن اشاره کرد و تصمیمات مدیریتی در این بخش از اهمیت بسیار بالایی برخوردار هستند. اتخاذ تصمیمات صحیح‌تر منوط به وجود داده‌های دقیق‌تر می‌باشد. در بحث مدیریت حمل و نقل از پر استفاده‌ترین ابزارها می‌توان به ماتریس سفر اشاره نمود. تهیه ماتریس سفر از روش‌های متداول مانند آمارگیری مبدا-مقصد انجام می‌شود که بسیار زمان‌بر و پرهزینه می‌باشد. لذا ضرورت دارد تا روش‌هایی ایجاد شوند که بتوانند با صرف زمان و هزینه بسیار کمی، ماتریس سفر را با دقت کافی و قابل قبولی تخمین بزنند و هرچه داده‌های مورد استفاده راحت‌تر جمع‌آوری شوند و همچنین روش تخمین با دقت بالاتر و هزینه کمتری همراه باشد آن روش تخمین مناسب‌تر خواهد بود.

۱-۶ نوآوری موضوع

در این تحقیق برای اولین بار از پارامترهای اقتصادی-اجتماعی شهر برای تخمین ماتریس سفر استفاده شد. روش مورد استفاده در این تحقیق، تخمین ماتریس سفر با استفاده از هوش مصنوعی و به کمک شبکه عصبی نرم افزار متلب بود.

۱-۷ فرضیات تحقیق

- شاخص‌های اقتصادی-اجتماعی بر رفتار سفر و جابجایی مردم اثر دارند.
- شاخص‌های اقتصادی-اجتماعی در سازمان‌ها موجود اند و تهیه آنها سریع‌تر و ارزان‌تر از آمارگیری ترافیکی است.
- بین تاثیر شاخص‌های اقتصادی-اجتماعی در الگوی ترافیکی کلانشهرها مشابهت مناسبی وجود دارد.

۸-۱ فرضیه های تحقیق

- با استفاده از هوش مصنوعی می توان یک ماتریس سفر را از شاخص های اقتصادی-اجتماعی تخمین زد.
- تخمین با هوش مصنوعی نسبت به دقت روش های سنتی دارای دقتی کافی و قابل قبول است.
- روش های مختلف استفاده از شبکه عصبی بر نتیجه تخمین سفر نتایج متفاوتی ارائه می دهند و باید روش مناسب تر را یافت.

۹-۱ ساختار پایان نامه

این تحقیق در شش فصل تنظیم شده است.

در فصل اول کلیات موضوع تحقیق شامل طرح و تعریف مسئله، ضرورت و هدف انجام مطالعه و در نهایت متدولوژی تحقیق تشریح شده است.

در فصل دوم ابتدا مبانی روش های تهیه ماتریس سفر بررسی شدند سپس پیرامون تخمین ماتریس سفر مطالبی بیان شد و پس از آن مدل های متداول تخمین ماتریس سفر بررسی شدند. و در نهایت مطالعات صورت گرفته در زمینه تخمین ماتریس سفر و همچنین مطالعاتی که در آنها با استفاده از هوش مصنوعی به تخمین پارامترهای ترافیکی پرداخته اند بررسی شدند.

در فصل سوم مبانی و توضیحاتی پیرامون شبکه های عصبی و روش های متداول تخمین با استفاده از هوش مصنوعی از جمله روش های شبکه عصبی، روش های فازی و سیستم های عصبی-فازی بیان شد.

در فصل چهارم به بررسی داده های مورد بررسی در این تحقیق و نحوه آماده سازی داده ها و بررسی ضریب همبستگی میان داده ها با استفاده از نرم افزار IBM SPSS Statistics 21 و MATLAB

2013 پرداخته شد. در نهایت نیز کد برنامه‌های مورد استفاده در نرم افزار MATLAB 2013 که

جهت انجام تحقیق استفاده شده اند شرح داده شدند.

در فصل پنجم داده‌های نهایی با استفاده از روش‌های مورد بررسی در این تحقیق جهت تخمین

استفاده شدند و نتایج تخمین‌ها در این فصل آورده شد. سپس پیرامون تحلیل نمودارها و

خروجی‌های نتایج بدست آمده مطالبی ذکر شد.

در نهایت و در فصل ششم جمع‌بندی و نتیجه‌گیری نتایج به دست آمده مطرح می‌شود و

همچنین پیشنهاداتی برای پژوهشگرانی که در آینده تصمیم بر ادامه دادن این تحقیق خواهند داشت

مطرح می‌گردد تا موضوع این تحقیق از جنبه‌های دیگری نیز مورد بررسی قرار گیرد.

فصل ۲ : پیشینه تحقیق

۲-۱ مقدمه

در این فصل به بررسی کارهای انجام شده در گذشته پرداخته می‌شود تا با استفاده از نتایج آنها بتوان مدل‌های بهتری را مورد استفاده قرار داد و نتیجه مطلوب‌تری حاصل شود.

ابتدا به بررسی روش‌های متداول تهیه ماتریس سفر از تهیه بوسیله آمارگیری تا روش‌های تخمینی پرداخته و سپس به معرفی اجمالی نحوه تخمین ماتریس سفر با استفاده از هوش مصنوعی پرداخته شده است. علیرغم تحقیق و بررسی منابع علمی متعدد تحقیقات چندانی مشابه تحقیق پیش رو یافت نشد و لذا به ذکر چند نمونه مقالات علمی‌ای که می‌توان از نتایج‌شان در این تحقیق بهره برد پرداخته خواهد شد و مقالاتی که در آنها پارامترهای ترافیکی با استفاده از هوش مصنوعی تخمین زده شده اند را بررسی خواهیم کرد تا علاوه بر پی‌بردن به الگوریتم‌های یادگیری متناسب با روش تحقیق، پی برده شود که آیا تخمین ماتریس سفر بوسیله هوش مصنوعی امری ممکن یا ناممکن است؟

۲-۲ معرفی موضوع

مسئله مورد نظر این است که بتوان با استفاده از شبکه عصبی مصنوعی که یک روش تقریباً جدید برای تخمین داده‌های آماری می‌باشد، بر اساس جمعیت ساکن در نواحی شهری و زیربنای کاربری اراضی موجود در هر ناحیه تخمین نسبتاً دقیقی از ماتریس سفر بین نواحی شهری ارائه نمود.

در مطالعات و برنامه‌ریزی حمل‌ونقل شهری یکی از مهم‌ترین اطلاعات مورد نیاز مدیران و مهندسين دانستن تعداد سفرهای انجام شده بین مراکز مختلف شهر مورد مطالعه است که از آن به عنوان اطلاعات تقاضای سفر یاد می‌شود. اطلاعات تقاضای سفر به صورت ماتریس‌های مبدأ-مقصد ارائه شده و هرچه درایه‌های ماتریس سفر از دقت بالاتری برخوردار باشند به همان میزان امکان برنامه‌ریزی و تصمیم‌گیری دقیق‌تری را در آینده فراهم می‌سازند. امروزه از مسائل مطرح در علم حمل و نقل بدست آوردن ماتریس‌های تقاضا و بهنگام‌سازی ماتریس‌های قبلی توسط کم‌هزینه‌ترین و

سریع‌ترین روش‌ها است. در گذشته از روش‌های متداولی برای تخمین ماتریس سفر در سال طرح با استفاده از ماتریس سفر سال‌های پیشین استفاده می‌شده است. از جمله روش‌های فراوانی گرایانه^۱ که در آن تنها به مشاهدات و بسامد رخدادها استناد می‌شود، از جمله تخمین ماتریس سفر با استفاده از اطلاعات حجم کمان (روش حداقل مربعات خطا، روش حداکثر جابجایی، روش بیشترین درست‌نمایی، روش حداقل اطلاعات) و تخمین ماتریس سفر با استفاده از ثبت پلاک خودروها و یا تخمین بیزین^۲ و یا روش‌های کمتر متداول دیگر.

در این تحقیق به تخمین ماتریس سفر با استفاده از روش جدیدی یعنی با استفاده از ابزار شبکه عصبی در هوش مصنوعی پرداخته شده است.

۳-۲ روش‌های تهیه ماتریس سفر

۳-۲-۱ آمارگیری

آمارگیری مبدأ-مقصد را می‌توان به روش‌های مختلفی انجام داد که با توجه به وقت‌گیر و پرهزینه بودن هر ۱۰ الی ۲۰ سال انجام می‌شود، حال در ادامه انواع آن معرفی می‌شود.^۳

• آمارگیری مشاهده‌ای

این روش به وسیله مشاهده مستقیم وضعیت ترافیک و برداشت اطلاعات لازم از آن، برنامه‌ریزی شده و انجام می‌گیرد. روش "شماره وسیله نقلیه" که در آن با ثبت شماره پلاک وسیله نقلیه، مسیر آن ردیابی می‌شود از این نوع روش‌ها است. روش‌های "برچسب بر وسایل نقلیه" و "چراغ روشن"، از دیگر روش‌هایی هستند که توسط آنها محل شروع و پایان سفر را می‌توان تعیین کرد، این روش‌ها نیز

^۱ Frequentist

^۲ Bayesian

^۳ Cambridge Systematics, inc., (1996) "Travel Survey Manual," A paper prepared for U.S. Department of Transportation & U.S. Environmental Protection Agency.

در گروه روش‌های مشاهده‌ای جای می‌گیرند. این روش معمولاً برای تعیین مبدأ-مقصد سفرهای واقع در یک شبکه ترافیکی کوچک کاربرد دارد و برای استخراج مبدأ-مقصد سفرهای یک شهر قابل کاربرد نیست.

- آمارگیری پرسشنامه‌ای

در این روش، اطلاعات به وسیله توزیع پرسشنامه گردآوری می‌شود. به دلیل حجم بالای اطلاعات مورد نیاز در پرسشنامه‌های مبدأ-مقصد خانوار، فقط در مطالعات جامع حمل‌ونقل از این روش استفاده می‌شود. روش پرسشنامه‌ای به پنج زیرگروه ذیل تقسیم می‌شود:

- مصاحبه حضوری
- مصاحبه تلفنی
- کارت پستی
- روش ترکیبی

۲-۳-۲ مدل‌های تخمینی

در این مدل‌ها از نمونه برداری و تخمین تعداد کل سفرهای انجام شده بوسیله ی روش های آماری با استفاده از داده های نمونه برداری شده استفاده می‌شود. به طور کلی هدف از نمونه برداری کسب اطلاعاتی درباره ویژگی افراد جامعه بدون صرف هزینه‌های زیاد است که بیانگر همان اطلاعات از همه افراد جامعه باشد. از این رو در نمونه‌برداری ابتدا می‌بایست نوع اطلاعاتی که قرار است ارزیابی شده و به جامعه نسبت داده شود، تعیین گردد. این اطلاعات با توجه به هدف کلی مطالعه و پارامترهای مورد نیاز برای مطالعه تعیین شده و اغلب به آن‌ها پارامترهای جامعه گفته می‌شود. استنباط‌های آماری را می‌توان راجع به میانگین این پارامترها انجام داد. برای نمونه درآمد و نرخ سفر خانوار، پارامترهای جامعه به حساب می‌آیند. برای تعیین پارامتر مناسب در یک مطالعه به تجربه کافی و اطلاع از نوع

کاربرد پارامتر در استنباط، تحلیل و یا مدل سازی نیاز است.

همان طور که می دانید اطلاعات تقاضای سفر به صورت ماتریس های مبدأ-مقصد یکی از اصلی ترین و ضروری ترین ورودی ها در برنامه ریزی حمل و نقل است و برای طراحی و مدیریت سیستم های حمل و نقل یکی از اطلاعات اساسی به حساب می آید. روش های استفاده از مدل، خود به سه دسته ذیل تقسیم می شوند.

✓ دسته اول: مدل های تقاضا

مانند مدل های تولید، جذب و توزیع سفر به ویژه مدل جاذبه. برای برآورد پارامترهای این مدل ها نیاز به اطلاعات سفرهای مبدأ-مقصد حداقل در یک مقطع زمانی می باشد.

✓ دسته دوم: مدل های برآورد ماتریس مبدأ-مقصد با استفاده از اطلاعات حجم ترافیک در

کمان های شبکه

✓ دسته سوم: مدل های آماری، اساس این مدل ها بر پایه نگرش آماری به مسئله است.

۲-۴ مدل های توزیع سفر

۲-۴-۱ مدل رشد:

در این روش برای کل منطقه مورد مطالعه یک ضریب عمومی رشد λ در نظر گرفته شده و ماتریس سفر سال پایه در آن ضرب می شود تا ماتریس سفر سال افق به دست آید (معادله ۲-۱). ضریب رشد بین دو ناحیه از میانگین ضرایب رشد آن ها به دست می آید.

$$T_{ij} = \lambda \times t_{ij}$$

معادله ۲-۱

t_{ij} : ماتریس سفر سال پایه

T_{ij} : ماتریس سفر سال افق طرح

λ : ضریب رشد سالیانه

۲-۴-۲ مدل دیترویت:

برخی از مدل‌های رشد، تنها تبادل سفرها را بدون در نظر داشتن جهت سفرها، در نظر می‌گیرند. یعنی مثلاً منطقه i بدون توجه به جذب یا تولید یک ضریب رشد λ_i دارد. از این نوع مدل‌ها می‌توان مدل‌های دیترویت و فراتار را نام برد. در روش دیترویت با استفاده از ضریب رشد هر ناحیه λ_i و رشد میانگین کل محدوده مورد مطالعه λ_{av} توزیع سفر در سال طرح به صورت زیر به دست می‌آید (معادله ۲-۲):

$$T_{ij} = t_{ij} \frac{\lambda_i \times \lambda_j}{\lambda_{ave}} \quad \text{معادله ۲-۲}$$

که در آن:

T_{ij} : ماتریس سفر سال افق طرح

T_{ij} : ماتریس سفر سال پایه

λ_i : ضریب رشد منطقه مبدا

λ_j : ضریب رشد منطقه مقصد

λ_{ave} : ضریب رشد میانگین مناطق مبدا و مقصد

۳-۴-۲ مدل جاذبه:

معروفترین و پرکاربردترین مدل توزیع سفر، مدل جاذبه است که برای به دست آوردن آن از مفهوم قانون جاذبه نیوتن استفاده شده است. در مدل جاذبه بدون استفاده مستقیم از ماتریس سفرهای سال پایه سفرهای سال طرح پیش‌بینی می‌شود. مدل جاذبه اولین بار برای سفرهای خرید بین شهری به صورت زیر به کار رفته است (معادله ۳-۲):

$$T_{ij} = \alpha O_i D_j c_{ij}$$

که در آن:

T_{ij} : تعداد سفر از مبدأ i به مقصد j

α : ضریب تناسب

O_i : تعداد سفر از مبدأ i

D_j : تعداد سفر به مقصد j

C_{ij} : تابع مقاومت سفر بین مبدأ و مقصد

۲-۵ مدل تخمین ماتریس سفر با استفاده از حجم کمان

در این روش با استفاده از اطلاعات حجم جریان در کمان‌های انتخاب شده از شبکه حمل و نقل شهری و ساختار شبکه، ماتریس تقاضای مبدأ-مقصد معابر منقطع توسط کمانهای انتخابی به دست می‌آید. به بیان ساده‌تر، هدف برآورد (یا تصحیح) ماتریس تقاضای مبدأ-مقصد بر اساس اطلاعات در دسترس شبکه (حجم جریان ترافیک در کمان‌های شبکه) است. بدیهی است پس از تخصیص ماتریس مبدأ-مقصد به دست آمده بر روی شبکه، انتظار می‌رود جریان در هر کمان تقریباً برابر جریان مشاهده شده شود. لازم به ذکر است که این برابری به معنای برابری مطلق ریاضی نبوده، منظور نزدیک بودن آماری دو مقدار به یکدیگر است.

در جدول ۲-۱ به طور خلاصه نمونه‌ای از تحقیقات انجام شده در زمینه مدل‌های تخمین ماتریس سفر ارائه گردیده است:

جدول ۱-۲ تحقیقات انجام شده در زمینه مدل های تخمین جدول سفر

نوع مدل	نام نویسنده	خلاصه مقاله
بیشترین درستنمایی	اسپایس [۱]	یک مدل ماکزیمم درستنمایی برای برآورد ماتریس های مبدأ-مقصد
	کسکتتا و نگوین [۲]	یک چارچوب کلی برای برآورد یا به روز کردن ماتریس های مبدأ-مقصد
	هازلتون [۳]	برآورد ماتریس های مبدأ-مقصد از جریان کمان در شبکه خلوت
بیشترین انترویی و کمترین اطلاعات	فیسک [۴]	ترکیب ماکزیمم آنترویی برآورد ماتریس سفر با تخصیص کاربر متعادل
	ویلسون [۵]	مدل آنترویی در شهر و مدل سازی منطقه ای
	چی و همکاران [۶]	برآوردگر ماکزیمم آنترویی و کمترین مربعات برای برآورد ماتریس مبدأ-مقصد
	ویلیام سن [۷]	برآورد ماتریس مبدأ-مقصد از شمارش ترافیک
	برینیر-گوت [۸]	برآورد ماتریس های مبدأ-مقصد با استفاده از فرمول بندی برنامه ریزی چند هدفه
	روبرت [۹]	مقایسه الگوریتم ها برای برآورد پارامترهای ماکزیمم آنترویی
	ونزوین و ویلیام سن [۱۰]	برآورد احتمالی ماتریس سفر از شمارش ترافیک
	یو گو [۱۱]	کشف تناقض در ترافیک شبکه با استفاده از برآورد ماکزیمم آنترویی
	بل [۱۲]	برآورد ماتریس های مبدأ-مقصد به وسیله کمترین مربعات
	بیرلیر و توینت [۱۳]	یک برآورد ماتریس مبدأ-مقصد که ساختار برجسته دارد
	کسکتتا [۱۴]	برآورد ماتریس های سفر و مطالعه داده ها: یک برآورد کمترین مربعات
کمترین مربعات	نای [۱۵]	استنتاج ماتریس های مبدأ-مقصد با یک برآوردگر جریان مسیر کمترین مربعات
	یانگ و همکاران [۱۶]	مسئله برآورد ماتریس مبدأ-مقصد کاربر متعادل
	جودی و همکاران [۱۷]	استفاده از روش های آماری در مسائل تئوری و عملی

از دهه هشتاد میلادی به این سو مدل های زیادی برای برآورد و بروز کردن ماتریس های تقاضا با استفاده از اطلاعات شمارش حجم در کمان ها پیشنهاد و استفاده شده است، که دقت آن ها وابسته به مدل مورد استفاده، داده های ورودی و مجموعه کمان های شمارش شده جهت انجام مدل سازی است. در این مدل سازی ها ممکن است از توابع هدف مختلفی (مانند؛ بیشترین آنترویی^۱، کمترین مربعات^۲ و

¹ Maximum Entropy

² Minimum Least Squares

بیشترین درست‌نمایی^۱، محدودیت‌های جریان گوناگون(مانند: تعادل استفاده‌کننده^۲، تخصیص تناسبی^۳) و ورودی‌های متفاوت(مانند: جذب و تولید سفر، حجم ترافیک شمارش شده، زمان سفر و یا ماتریس سفر تاریخی) استفاده شود.

۲-۵-۱ مدل کمترین مربعات

یکی از عمومی‌ترین و رایج‌ترین روش‌هایی که تاکنون استفاده شده روش کمترین مربعات می‌باشد که در بسیاری از مدل‌ها مورد استفاده قرار گرفته و رابطه آن از قرار ذیل می‌باشد(معادله ۲-۴):

$$F_2(T, \hat{T}) = \sum (T_{ij} - \hat{T}_{ij})^2 \quad \text{معادله ۲-۴}$$

این امکان وجود دارد که تعیین وزن مخصوصی برای مسائل تک‌هدفی لازم باشد. یکی از روش‌های انتخاب وزن‌دهی معیار راستی‌آزمایی و قابلیت اعتماد داده‌های مشاهده شده می‌باشد. بنابراین در این روش از واریانس بین داده‌ها استفاده کرده و فرمول کمترین مربعات به صورت ذیل خواهد شد(معادله ۲-۵):

$$F_2(T, \hat{T}) = \sum \frac{1}{\sigma_a^2} (T_{ij} - \hat{T}_{ij})^2 \quad \text{معادله ۲-۵}$$

از این روش نیز در مدل‌های مختلفی توسط افرادی از جمله: کری(۱۹۸۱)، کاسکتا(۱۹۸۴)، مک نیل و هندریکسون(۱۹۸۵)، اسپایس(۱۹۹۰)، بیرلایر و توینت(۱۹۹۵)، فلورین و چن(۱۹۹۵)، ماهر و ژانگ(۱۹۹۹)، بیانسو(۲۰۰۱)، کاسکتا و پوستورینو(۲۰۰۱)، ماهر(۲۰۰۱)، کودینا و بارسلو(۲۰۰۴)، دوبلاس و بنیتز(۲۰۰۵) و نیه(۲۰۰۵) استفاده شده است.

³ Maximum Likelihood

⁴ User Equilibrium

⁵ Proportional Assignment

۲-۵-۲ روش بیشترین آنتروپی (بیشترین جابجایی)

برای تشریح روش بیشترین آنتروپی، سیستمی متشکل از تعداد زیادی عناصر مجزا را در نظر بگیرید. توصیف کامل از چنین سیستمی نیازمند بررسی جزئیات کامل ریزحالت^۱های آن سیستم (در حد کوچکترین اجزاء قابل تفکیک) است. برای مثالی مرتبط با برنامه ریزی حمل و نقل، بررسی سیستم در حد ریزحالت شامل شناسایی هر فرد سفرکننده، مبدأ و مقصد سفر، طریقه سفر، زمان و سایر خصوصیات سفر او است.

در بسیاری از موارد کاربردی، کافی است که به سیستم با دیدی کلی تر و عمومی تر و در حد جزئیات میان حالت نگریسته شود. در ادامه مثال قبل، بررسی در حد میان حالت^۲ ممکن است تنها به صورت تعداد سفر بین هر مبدأ-مقصد در نظر گرفته شود. به طور کلی ممکن است ریزحالت های زیاد و متفاوتی، میان حالت یکسانی را ایجاد کنند. به عنوان مثال دو فرد که در یک ناحیه ترافیکی ساکن هستند ممکن است مقاصد سفر خود را با یکدیگر عوض کنند. در این حالت در ریزحالت سیستم تغییر به وجود می آید ولی شرایط میان حالت حفظ می گردد. غالباً سطح بالاتری نیز برای تعمیم و کلی نگری به سیستم وجود دارد که به آن کلان حالت^۳ گفته می شود. به عنوان مثال، تعداد کل سفرهایی که از یک کمان خاص در شبکه عبور می کنند و با کل تعداد سفرهایی که در یک ناحیه تولید و یا به آن جذب می شوند.

به منظور دستیابی به شاخص های قابل اعتماد در بررسی فعالیت سفرکنندگان، اغلب ساده تر است که مشاهدات در سطح کلان حالت ها که از درجه زیادی از کلی نگری برخوردار هستند به انجام رسد. در حقیقت اکثر اطلاعات موجود از سیستم جابجایی ها در این سطح قرار دارد. به این طریق به علت افزایش عدم قطعیت در پیش بینی های در سطح جزئی، برآوردها در مورد آینده معمولاً به

¹ Micro State

² Meso State

³ Macro State

توصیفات در حد کلان حالت محدود می‌شود. به عنوان مثال، پیش‌بینی جمعیت در یک ناحیه ساده‌تر از پیش‌بینی تعداد خانوارهای ساکن در آن ناحیه است که از یک طبقه خاص هستند.

مبنای روش بیشترین آنتروپی، پذیرفتن این قضیه است که احتمال وقوع تمامی ریزحالت‌های سازگار با کلان‌حالت‌های در نظر گرفته شده در مسئله یکسان است، مگر اینکه اطلاعاتی بر خلاف این امر وجود داشته باشد. در حقیقت، این یک فرض عقلانی است که به‌وسیله آن از جزئی‌نگری در سطح ریزحالت و میان‌حالت صرف‌نظر می‌شود. یک راه حل مناسب برای اعمال سازگاری ریزحالت‌ها و میان‌حالت‌ها با اطلاعات موجود در مورد کلان‌حالت‌ها، بیان اطلاعات به شکل قیدهای برابری در یک برنامه ریاضی است [۱۸]. از آنجا که در فرآیند برنامه‌ریزی حمل و نقل نوعاً به توصیف سیستم در حد میان‌حالت (فرضاً تعداد سفر مبادله شده بین مبدأ-مقصد) نیاز است، آن دسته از میان‌حالت‌هایی که دارای بیشترین سطح احتمال هستند (به واقعیت نزدیک‌تر هستند) با در نظرگیری قیدهایی در مورد کلان‌حالت‌ها (فرضاً تردد در سطح معابر شبکه) شناسایی می‌شوند. ویلسون در سال ۱۹۷۰ [۱۹] به معرفی مفهوم و مدل آنتروپی پرداخته است و بیشترین جابجایی داده‌های موجود را آنالیز کرده تا به توزیع احتمالی یکسانی برسد. وی رابطه ذیل (معادله ۲-۶) را برای این مدل در نظر گرفت:

$$W = \frac{TN!}{\prod_{ij} T_{ij}!} \quad \text{معادله ۲-۶}$$

که در این رابطه:

TN: تعداد کل سفرها

t_w : سفرهای انجام شده بین هر مبدأ و مقصد

برای سهولت می‌توان بجای استفاده از W از تابعی یکنواخت مانند $\log W$ استفاده نمود

(معادله ۲-۷). لذا می‌توان نوشت:

$$\log W = \log \frac{T!}{\prod_{ij} T_{ij}!} = \log T! - \sum_{ij} \log T_{ij}! \quad \text{معادله ۲-۷}$$

طبق تقریب استرلینگ می‌توان نوشت (معادله ۲-۸):

$$\ln X! = X \ln X - X \quad \text{معادله ۸-۲}$$

با استفاده از این تقریب، بهینه‌سازی با سهولت بیشتری انجام خواهد شد (معادله ۹-۲). بنابراین:

$$\log W = \log T! - \sum_{ij} (T_{ij} \log T_{ij} - T_{ij}) \quad \text{معادله ۹-۲}$$

جمله $\log T!$ یک عدد ثابت است و معمولاً می‌توان آن را از مسئله بهینه‌سازی حذف نمود. مابقی

معادله فوق به صورت رابطه ذیل نوشته می‌شود و به عنوان تابع آنتروپی شناخته می‌شود (معادله ۲-۲).

(۱۰).

$$F_1(T, t) = - \sum T_{ij} (\log T_{ij} - 1) \quad \text{معادله ۱۰-۲}$$

در صورتی که اطلاعات سفر پیشین در اختیار باشد رابطه ۸-۲ به صورت ذیل (معادله ۱۱-۲)

بازنگری می‌شود:

$$F_1(T, t) = - \sum T_{ij} (\log T_{ij}/t_{ij}) - T_{ij} + t_{ij} \quad \text{معادله ۱۱-۲}$$

که در آن:

F: تابع هدف

T_{ij} : کل سفر آینده

t_{ij} : کل سفر پیشین

۳-۵-۲ مدل بیشترین درست‌نمایی (maximum likely hood)

این روش یکی از قدیمی‌ترین و پر اهمیت ترین روش‌ها در نظریه برآورد است. فرض کنید I

مجموعه‌ای از مبداهای، j، مجموعه‌ای از مقصدها و $R=I..J$ مجموعه‌ای از جفت‌های مبدأ-مقصد باشد.

برای هر جفت مبدأ-مقصد، ماتریس هدف $\hat{g}$ به وسیله مشاهدات یک فرآیند پواسون با میانگین

به دست آمده است که احتمال مشاهده به صورت زیر است (معادله ۱۲-۲):

$$prob = \frac{(\rho_r g_r)^{\hat{g}_r} \times e^{-\rho_r g_r}}{\hat{g}_r!} \quad \text{معادله ۱۲-۲}$$

اکنون از طرفن لگاریتم می‌گیریم، چون لگاریتم یک تابع صعودی است، پس ماکزیمم کردن این رابطه با ماکزیمم کردن لگاریتم آن برابر است (معادله ۲-۱۳).

$$\max \sum (\widehat{g}_r \log(\rho_r g_r) - \rho_r g_r - \log \widehat{g}_r!) \quad \text{معادله ۲-۱۳}$$

با صرف‌نظر کردن از مقادیر ثابت، می‌توان مسئله ماکزیمم احتمال را به صورت زیر فرمول‌بندی کرد (معادله ۲-۱۴):

$$F_1(g, \widehat{g}) = -\sum (\rho_i g_{ij} - g_{ij} \log \widehat{g}_{ij}) \quad \text{معادله ۲-۱۴}$$

که در آن:

$\widehat{g}_{ij}$: کل سفر پیشین

F : تابع هدف

ρ_i : ضریب پواسون

g_{ij} : کل سفر آینده

۲-۵-۴ روش حداقل اطلاعات (minimum information)

هنگامی که اطلاعات و پیش‌نیازهای کاملی از قبیل حجم کمان‌های شبکه، زمان سفر هر کمان و... برای تخمین ماتریس سفر در دسترس باشد می‌توان ماتریس سفر نهایی قطعی را پیش‌بینی کرد. در بعضی مواقع که هیچ اطلاعات اولیه برای تخمین در دسترس نمی‌باشد پیش‌بینی و تخمین ماتریس سفر با خطای زیادی همراه خواهد بود. هدف از تخمین این است که بتوان با کمترین اطلاعات لازم ماتریس سفر را با خطای قابل قبول تخمین زد که از لحاظ اقتصادی نیز مقرون به صرفه خواهد بود. لذا به مدلی که ماتریس از این روش به دست می‌آید مدل کمترین اطلاعات گفته می‌شود.

روش بیشترین آنتروپی از قاعده کمترین اطلاعات گرفته شده است که در سال ۱۹۷۸ توسط ون

زویلن به صورت معادله ۲-۱۵ بیان شد:

$$F_1(g, \hat{g}) = \sum g_{ij} (\log \frac{g_{ij}}{\hat{g}_{ij}} - 1)$$

که در آن

F : تابع هدف

g_{ij} : کل سفر آینده

$\hat{g}_{ij}$: کل سفر پیشین

در رابطه فوق اگر ماتریس اولیه‌ای در دسترس نباشد تمامی مقادیر آن را یکسان در نظر گرفته و یا با در نظر گرفتن وزن یکسان برای ماتریس اولیه رابطه فوق به رابطه آنتروپی تبدیل خواهد شد.

۲-۶ مقاله Hichem Omrani [۲۰]

پس از تحقیق فراوان در منابع معتبر علمی و نشریات بین المللی، نتیجه حاصل شد که تا کنون مطالعات مشابهی در زمینه برآورد ماتریس سفر با استفاده از هوش مصنوعی بر اساس طبقه‌بندی‌های جمعیتی و سطح زیربنای کاربری اراضی واقع در نواحی مختلف شهری انجام نشده است. لذا در ادامه به بررسی مقالات معدودی که می‌توان از نتایج آنها در موضوع تحقیق بهره برد پرداخته شد:

• روش تحقیق

در مقاله omrani، چهار روش یادگیری ماشین

artificial neural net-MLP, artificial neural net-RBF, multinomial logistic regression, and support vector machines

برای پیش بینی روش سفر افراد در شهر لوکزامبورگ استفاده شد. در روش های ارائه شده از ویژگی های افراد^۱، مشخصات روش حمل و نقل^۲ و داده های مربوط به محل کار و محل سکونت استفاده می‌شود. داده ها از آمارگیری ملی PSELL survey (Socio-economic Panel Survey

¹ individuals' characteristics

² transport mode specifications

(Liewen zu L'etzebuerg) بدست آمد که این آمارگیری مربوط به سال ۲۰۰۳ است و در آن ۳۶۷۰ خانوار یعنی ۹۵۰۰ نفر آمارگیری شدند که برآورد شاخص های اجتماعی، جمعیت شناختی و اقتصادی را برای کل جمعیت امکان پذیر می کند و شامل اطلاعاتی در مورد جابجایی روزانه (به عنوان مثال ، از خانه تا محل کار) افرادی است که یا در لوکزامبورگ زندگی می کنند یا کار می کنند. ویژگی های فردی برای بررسی ارتباط جابجایی های روزانه (به ویژه سفرهای بین خانه و محل کار) با ویژگی های افراد شاغل استخراج شدند. همچنین از اطلاعات مربوط به حمل و نقل عمومی و موقعیت جغرافیایی محل های مسکونی و کاری استفاده شد. در نهایت نیز نرخ پیش بینی موفق به دست آمده توسط شبکه های عصبی و چندین روش جایگزین برای پیش بینی انتخاب حالت سفر با استفاده از RMSE (root MSE) و average probability of correct assessment (APCA) مقایسه شدند.

داده های مربوط به آمارگیری نشان دادند که ۸۳٪ از کارگران ساکن در لوکزامبورگ از خودروهای شخصی خود برای رفت و آمد به محل کار استفاده می کنند. نسبت افرادی که از وسایل نقلیه عمومی استفاده می کنند فقط ۱۱٪ بود. در حمل و نقل عمومی، اتوبوس حالت غالب (۸۰٪ از مسافران) بود. و ۶٪ از کارگران از دوچرخه استفاده می کردند.

• پارامترهای مورد بررسی

هزینه سفر، درآمد، سن، جنسیت، ملیت، نوع خانواده، تحصیلات، مالکیت خودرو، تعداد ایستگاه های اتوبوس و ایستگاه های قطار در منطقه محل سکونت، منطقه محل سکونت و منطقه محل کار.

• نتایج تحقیق

نتایج نشان داد که شبکه های عصبی مصنوعی (ANN) در مقایسه با سایر جایگزین ها عملکرد بهتری دارند (جدول ۲-۲). با استفاده از معیار کارایی APCA (میانگین احتمال ارزیابی صحیح) SVM و MNL نیز مناسب بنظر می رسند. با این حال ANN عملکرد کمی بهتری داشت که دقت پیش بینی را افزایش می دهد.

جدول ۲-۲ مقایسه نتایج

	APCA×100	APCA×100	APCA×100	APCA×100
	Overall	(Car)	(PT)	(Soft mode)
MNL	64.68 ± 1.16	78.48 ± 0.14	22.68 ± 0.95	7.29 ± 0.50
SVM	67.70 ± 3.81	79.96 ± 2.56	23.01 ± 0.83	26.49 ± 7.16
ANN-RBF	79.24 ± 4.54	80.96 ± 2.83	35.01 ± 6.71	31.27 ± 8.71
ANN-MLP	81.12 ± 1.12	85.10 ± 1.23	40.65 ± 2.27	25.63 ± 9.84

۷-۲ مقاله Gusri Yaldi [۲۱]

• روش تحقیق

Remya & Mathew از شبکه‌های عصبی مصنوعی برای تخمین ماتریس O-D استفاده کردند. از داده‌های مربوط به سفر کاری براساس آمارگیری مصاحبه ای در سال ۲۰۰۵ در شهر پادانگ، اندونزی استفاده شد. داده ها به سه بخش تقسیم شدند (تعلیم، ارزیابی، تست) و از آن جهت تخمین ماتریس سفر (تعداد سفرها و میزان تولید و جذب هر ناحیه) استفاده شد.

در این مطالعه برای انتخاب کوتاه‌ترین مسیر از الگوریتم Dijkstra و برای آموزش هوش مصنوعی از الگوریتم های Levenberg-Marquardt (LM) و Back propagation (BP) و Variable Learning Rate (VLR) استفاده شد تا تاثیر الگوریتم آموزش بر نتایج حاصل از تخمین مشخص شود. و همچنین برای ارزیابی عملکرد هوش مصنوعی از تکنیک Root Mean Square Error (RMSE) استفاده شد. با مشاهده مقادیر RMSE و R^2 و واریانس داده های برآورد شده و مقایسه آنها با مقادیر مشاهده شده نتیجه گرفته شد که هوش مصنوعی در صورت انتخاب الگوریتم مناسب جهت تعلیم می تواند در تخمین ماتریس سفر به خوبی عمل کند و از میان سه الگوریتم مذکور جهت تعلیم هوش مصنوعی، الگوریتم LM بهترین نتایج را ارائه می دهد

داده ها مربوط به سفر کاری و براساس آمارگیری مصاحبه ای در سال ۲۰۰۵ در شهر پادانگ، اندونزی گردآوری شد. این منطقه شامل ۳۶ منطقه است.

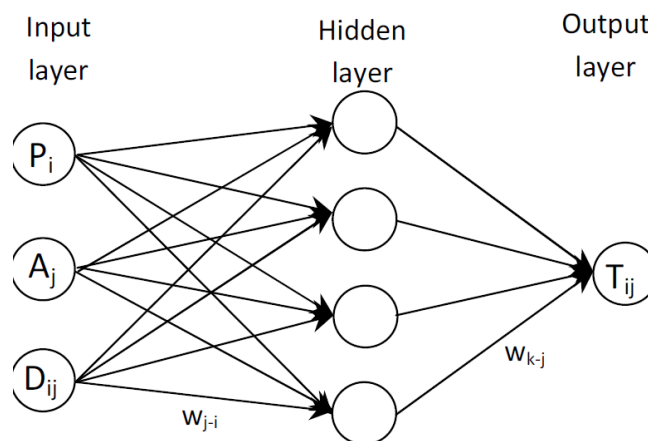
سپس داده ها برای تعلیم شبکه عصبی به سه بخش مطابق جدول ۲-۳ تقسیم شدند :

جدول ۲-۳: بلوک بندی داده ها

Block number	Zone numbers	Total (Zones)	Remarks
1 st Block, Training zones	20,26,33,6,3,9,32,16,22,1,12,8,19,35	14	40% of total zone numbers
2 nd Block, Validation zones	29,25,23,36,10,4,14,15,27,30,34	11	30% of total zone numbers
3 rd Block, Testing zones	21,11,31,17,18,2,13,28,24,5,7	11	30% of total zone numbers
Total		36	100%

• معماری شبکه عصبی :

این شکل سه گره ورودی را نشان می دهد تولید سفر (P_i) ، جذب سفر (A_j) و فاصله (D_{ij}). در لایه خروجی یک گره وجود دارد به نام تعداد سفر تخمینی (T_{ij}). هر گره با وزن اتصال W_{kj} و W_{ji} به گره های لایه پنهان متصل می شود (شکل ۲-۱).



شکل ۲-۱: معماری شبکه عصبی در این تحقیق

در جدول ۴-۲ می‌توان نتایج بدست آمده را در ۳۰ مرحله آزمایش مشاهده کرد

جدول ۴-۲ : عملکرد آزمایش شبکه عصبی آموزش داده شده با الگوریتم های متفاوت

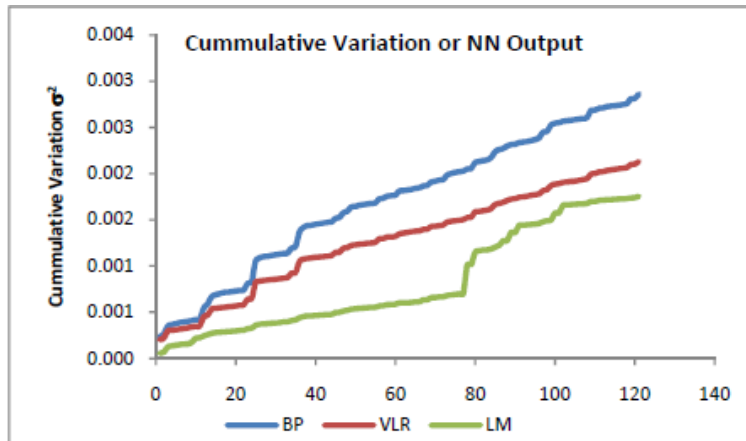
Experiment #	BP		VLR		LM	
	RMSE	R ²	RMSE	R ²	RMSE	R ²
1	178	0.016	161	0.251	136	0.404
2	187	0.000	158	0.243	145	0.330
3	195	0.014	195	0.015	129	0.505
4	201	0.052	162	0.215	132	0.542
5	208	0.056	156	0.249	158	0.246
6	202	0.050	166	0.197	126	0.502
7	172	0.052	130	0.507	128	0.468
8	161	0.249	161	0.249	131	0.445
9	165	0.176	165	0.176	134	0.499
10	174	0.080	174	0.088	134	0.464
11	141	0.376	141	0.376	130	0.480
12	186	0.098	189	0.150	132	0.446
13	142	0.350	140	0.361	133	0.452
14	169	0.107	154	0.239	138	0.403
15	166	0.104	166	0.103	132	0.444
16	192	0.042	159	0.268	145	0.344
17	172	0.029	173	0.027	144	0.426
18	182	0.009	167	0.116	127	0.474
19	181	0.000	151	0.330	136	0.425
20	175	0.027	175	0.026	134	0.425
21	179	0.163	179	0.163	125	0.501
22	146	0.358	155	0.343	132	0.445
23	200	0.031	146	0.318	143	0.483
24	200	0.051	156	0.235	141	0.369
25	183	0.021	152	0.293	138	0.392
26	213	0.023	174	0.022	125	0.507
27	173	0.023	147	0.310	136	0.493
28	212	0.025	156	0.267	169	0.143
29	195	0.080	175	0.017	127	0.506
30	171	0.084	191	0.097	134	0.437
Average1	181	0.092	162	0.208	136	0.433
Average2	168	0.194	152	0.315	125	0.505

با بررسی مقادیر RMSE و R² می‌توان متوجه شد که نتایج تخمین با هوش مصنوعی به هنگام

استفاده از الگوریتم LM در تعلیم هوش مصنوعی، بصورت میانگین مقادیر RMSE کمتری و R²

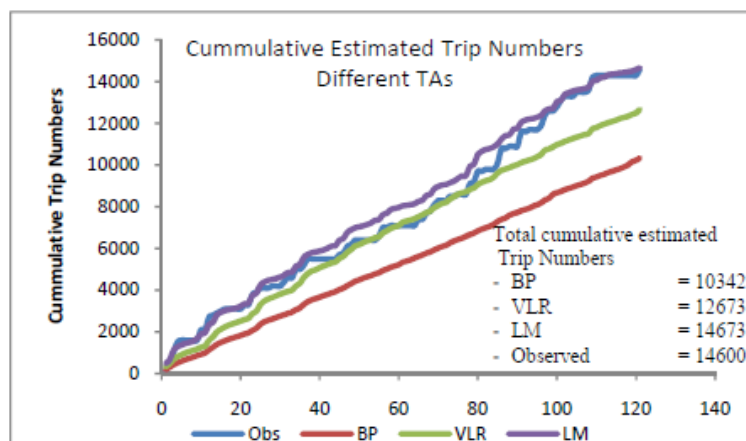
بیشتری در مقایسه با الگوریتم های آموزش دیگر مشاهده شد که دلیلی بر تخمین نزدیک به واقعیت

تر توسط هوش مصنوعی در هنگام استفاده از الگوریتم LM در این مطالعه، می‌باشد.



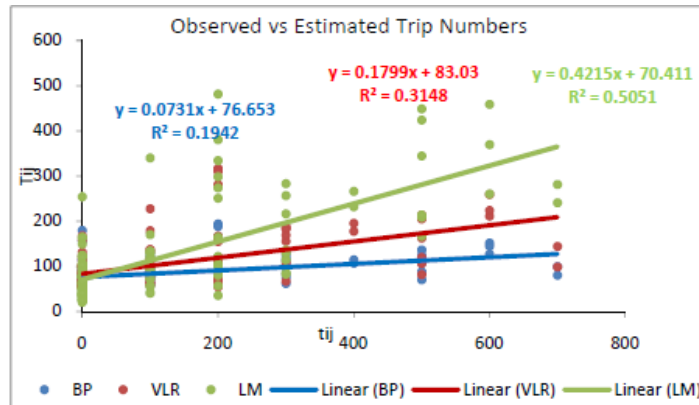
شکل ۲-۲: واریانس نتایج

شکل ۲-۲ واریانس میان نتایج را نشان می‌دهد و با توجه به نمودار بدست آمده مشاهده می‌شود نتایج به هنگام استفاده از الگوریتم LM واریانس کمتری نسبت به دیگر حالات، دارد.



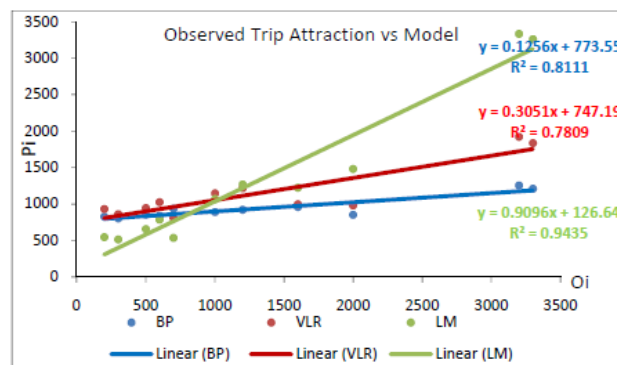
شکل ۳-۲: تعداد تجمعی سفرهای برآورد شده

شکل ۳-۲ تعداد تجمعی سفرهای برآورد شده توسط هوش مصنوعی را نشان می‌دهد و مشاهده می‌شود آموزش مدل NN با الگوریتم BP تعداد سفرهای بسیار پایین تری از مقدار واقعی ارائه می‌دهد. اما استفاده از الگوریتم LM مقادیری بسیار مشابه با مقادیر مشاهده شده، برآورد می‌کند.



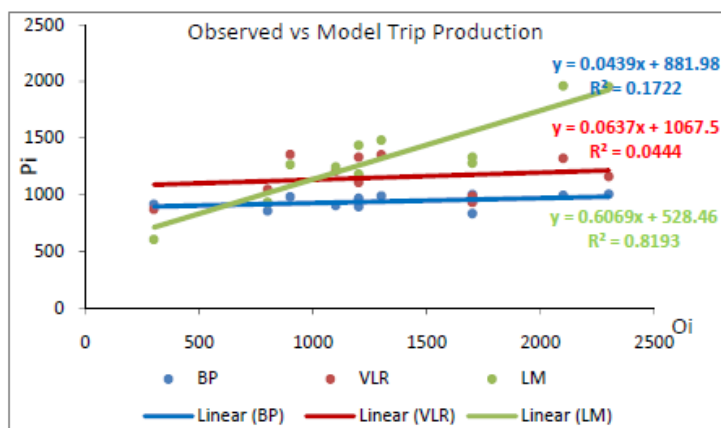
شکل ۲-۴: رابطه رگرسیون تعداد سفرهای تخمین زده شده

شکل ۲-۴ رابطه بین تعداد سفرهای مشاهده شده و تخمین زده شده را نشان می‌دهد و با توجه به آنکه شیب خط متناظر با نتایج تخمین هنگام استفاده از الگوریتم LM با خط متناظر با زاویه ۴۵ درجه نزدیکتر است، لذا نتایج بدست آمده در هنگام استفاده از این الگوریتم، قابل قبول تر از نتایج دیگر الگوریتم‌ها هستند.



شکل ۲-۵: رابطه رگرسیون برآورد سفرهای جذب شده

شکل ۲-۵ رابطه میان مقادیر جذب سفر تخمین زده شده توسط هوش مصنوعی را در برابر مقادیر تولید مشاهده شده نشان می‌دهد و با توجه به آنکه شیب خط متناظر با نتایج تخمین هنگام استفاده از الگوریتم LM با خط متناظر با زاویه ۴۵ درجه نزدیکتر است، لذا نتایج بدست آمده در هنگام استفاده از این الگوریتم، قابل قبول تر از نتایج دیگر الگوریتم‌ها هستند.



شکل ۲-۶: رابطه رگرسیون برآورد سفرهای تولید شده

شکل ۲-۶ رابطه میان مقادیر تولید سفر تخمین زده شده توسط هوش مصنوعی را در برابر مقادیر تولید مشاهده شده نشان می‌دهد و با توجه به آنکه شیب خط متناظر با نتایج تخمین هنگام استفاده از الگوریتم LM با خط متناظر با زاویه ۴۵ درجه نزدیکتر است، لذا نتایج بدست آمده در هنگام استفاده از این الگوریتم، قابل قبول تر از نتایج دیگر الگوریتم‌ها هستند. همچنین مقدار R^2 بسیار مناسب تری (برابر با ۰/۸) نسبت به باقی حالات دارد.

• نتایج تحقیق

از نتایج فوق می‌توان نتیجه گرفت، شبکه عصبی می‌تواند در صورت انتخاب الگوریتم آموزش صحیح و معماری ساختاری مناسب و متناسب با داده های ورودی، نتایج قابل قبولی در برآورد ماتریس سفر ارائه دهد

در میان الگوریتم های مورد آزمایش جهت تعلیم شبکه عصبی ، استفاده از الگوریتم Levenberg Marquardt algorithm منتج به دستیابی به نتایج بهتر و دقیق تر در برآورد پارامتر های ماتریس سفر شد

٢-٨ مقاله Albertengo [٢٢]

● روش تحقیق

در این مقاله داده های برداشت شده از سنسورهای سطح راه شهر torin را ابتدا پیش پردازش کرده و سپس آنها را به یک هوش مصنوعی وارد کرده و جهت تخمین ماتریس سفر، تولید و جذب تعداد سفر نواحی از آن استفاده شد

ابتدا از سایت opendata.5t.torino.it/get_fdt داده های ترافیکی برداشت شد که این اطلاعات حاوی تعداد و سرعت متوسط وسایل نقلیه در منطقه شهری تورین (ایتالیا) است که بوسیله سنسور های سطح راه که موقعیت آنها در شکل ۲-۷ درج شده است برداشت شده اند.



شکل ۲-۷: موقعیت سنسورهای برداشت اطلاعات ترافیکی در سطح شهر تورن

داده ها مربوط به بازه زمانی 01-Oct-2017 تا 28-Feb-2018 می باشند و شامل ۱۰۳ روز هفته،

۲۰ شنبه، ۲۱ یکشنبه و ۷ تعطیلات می باشند.

تجزیه و تحلیل داده ها نشان می دهد که داده ها را می توان به چهار نوع روز طبقه بندی کرد

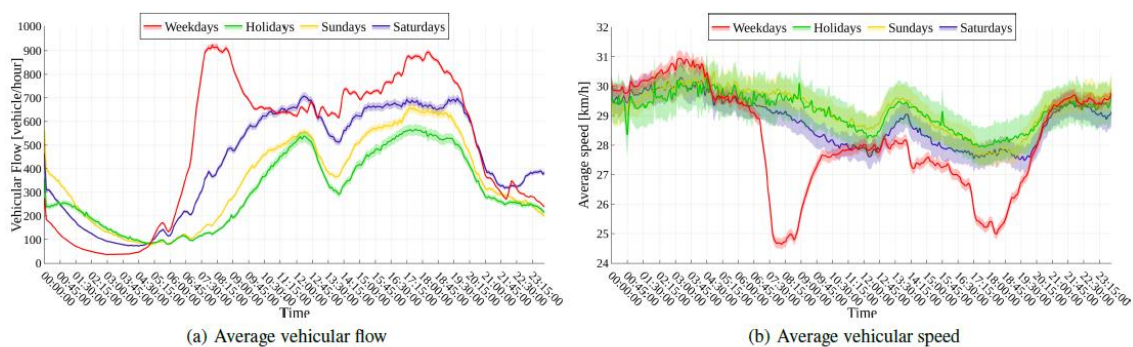
(Fox and Clark, ۱۹۹۸)

۱. روزهای هفته (از دوشنبه تا جمعه)

۲. شنبه ها

۳. یکشنبه ها

۴. تعطیلات (تعطیلات ملی و محلی)



شکل ۲-۸: جریان و سرعت ترافیک در شهر تورین بر اساس نوع روز

• جریان :

در طول روزهای هفته (شکل ۲-۸)، دو قله بلند از ۰۷:۴۰ تا ۸:۵۰ و از ۱۷:۱۰ تا ۱۸:۳۵ است. این قله ها به دلیل مسافرانی است که بین خانه و محل کار و بازگشت به آنجا مسافرت می کنند. شنبه ها و یکشنبه ها نسبت به روزهای هفته ترافیک کمتری را تجربه می کنند و بالاترین قله ها بین ۱۲:۰۵ تا ۱۲:۵۰ است.

• سرعت :

با توجه به سرعت وسایل نقلیه در چهار دسته بندی روزانه مختلف در این تحقیق (شنبه ها، یکشنبه ها و تعطیلات) (شکل ۲-۸) سرعت وسایل نقلیه روندی مشابه را دنبال می کنند و میانگین ماندگاری آن در حدود ۲۹ کیلومتر در ساعت است. در روزهای هفته ترافیک در ساعات اوج کاهش می یابد

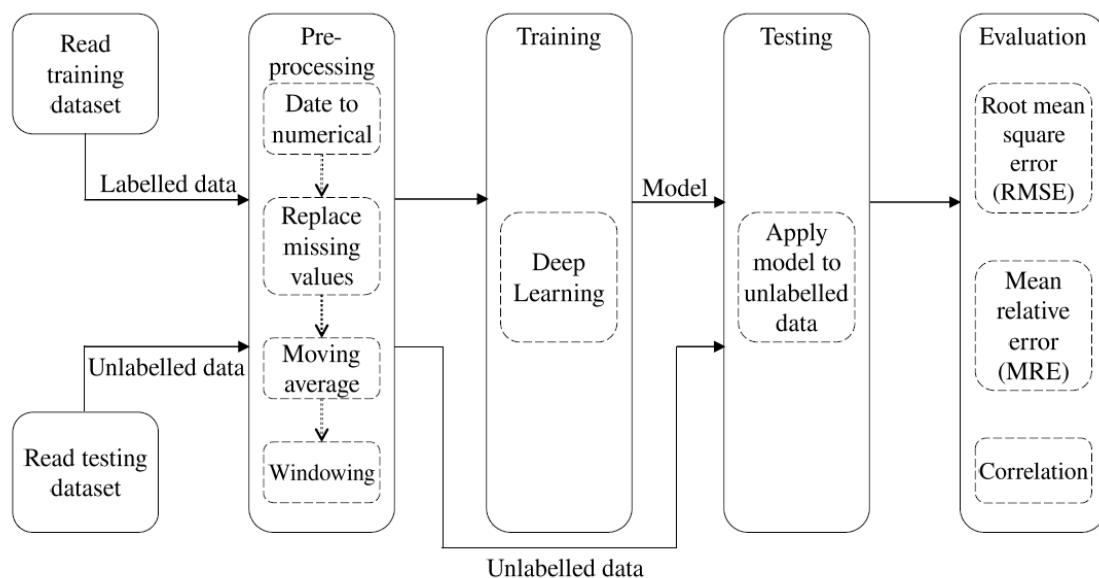
(یعنی از ۰۷:۲۰ تا ۰۹:۵۵ و از ۱۶:۵۵ تا ۱۹:۲۵). این به دلیل جریان زیاد وسایل نقلیه در آن دوره ها است که باعث ایجاد انسداد می شود.

- متدلوژی تحقیق :

برای پیش بینی شدت ترافیک از تحلیل رگرسیون پنجره‌ای نظارت شده با یادگیری عمیق (supervised windowbased regression analysis with Deep Learning) استفاده شد.

این فرآیند مطابق فلوچارت زیر انجام شد، بدین صورت که ابتدا داده ها نامگذاری شده و بر روی آنها پیش پردازش انجام گرفته و بصورت یادگیری عمیق به هوش مصنوعی درصد مشخصی از داده ها را وارد کرده و بوسیله آنها هوش مصنوعی تعلیم داده می شود. پس از آن هوش مصنوعی داده ها را پردازش کرده و پارامترهای مورد انتظار (تعداد سفر بین هر دو ناحیه و مقادیر تولید و جذب نواحی) را برآورد می کند.

سپس داده ها با مقادیر مشاهده شده مقایسه شده و عمل ارزیابی هوش مصنوعی انجام می پذیرد پس از آن بوسیله تکنیک های RMSE و MRE داده های تخمین زده شده با داده های واقعی مقایسه می شوند.



شکل ۲-۹ : فلوچارت پردازش داده ها بوسیله هوش مصنوعی

در فلوچارت پردازش داده ها (شکل ۲-۹) بخش های زیر را در بخش Pre-processing مشاهده

می کنیم :

- Date to numerical:

در این بخش برای مثال داده های روز دوشنبه ساعت ۰۰:۰۰ را با رقم ۰ و ساعت ۰۰:۰۵ را با رقم

۵ نشان داده و به همین صورت مابقی داده ها اصلاح شدند

- Replace missing values:

بوسیله درونیایی خطی داده های از دست رفته جایگزین شدند

- Time series windowing:

برای مثال مطابق شکل ۲-۱۰ داده ها بصورت پنجره های سری زمانی در می آیند تا نتایج بهتری

از پردازش حاصل گردد (هر پنجره برابر با ۳۰ دقیقه و فواصل پنجره ها ۵ دقیقه).

Original time series

Windowed cross-sectional format

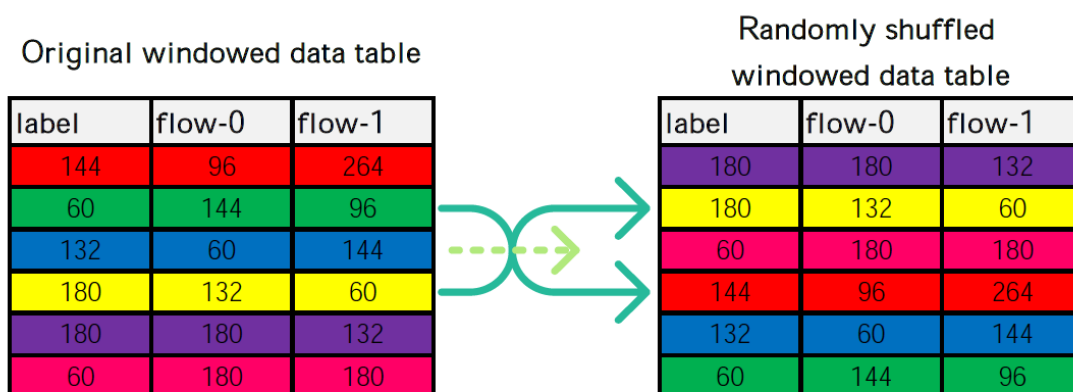
flow	label	flow-0	flow-1	flow-2	flow-3	flow-4	flow-5
12	144	96	264	108	192	156	12
156	60	144	96	264	108	192	156
192	132	60	144	96	264	108	192
108	180	132	60	144	96	264	108
264	180	180	132	60	144	96	264
96	60	180	180	132	60	144	96
144	84	60	180	180	132	60	144
60	108	84	60	180	180	132	60
132	132	108	84	60	180	180	132
180	96	132	108	84	60	180	180
180	60	96	132	108	84	60	180

شکل ۲-۱۰: مثالی از پیش پردازش پنجره های سری زمانی

- Shuffled order:

این بخش از پیش پردازش بصورت تصادفی داده‌ها را جابجا می‌کند (شکل ۲-۱۱) تا تعیین شود که با تغییر موقعیت داده‌ها چه تغییری در نتایج ایجاد می‌شود.

برای مثال بصورت زیر پس از پنجره سازی سری زمانی، جای ردیف‌ها بصورت رندم تغییر داده می‌شود.



شکل ۲-۱۱: مثالی از جابجایی تصادفی جدول داده‌های پنجره‌ای

- Moving average:

میانگین متحرک معمولاً با داده‌های سری زمانی برای هموار کردن نوسانات کوتاه مدت و برجسته سازی روند یا چرخه‌های طولانی مدت استفاده می‌شود.

- نرم افزار پردازش :

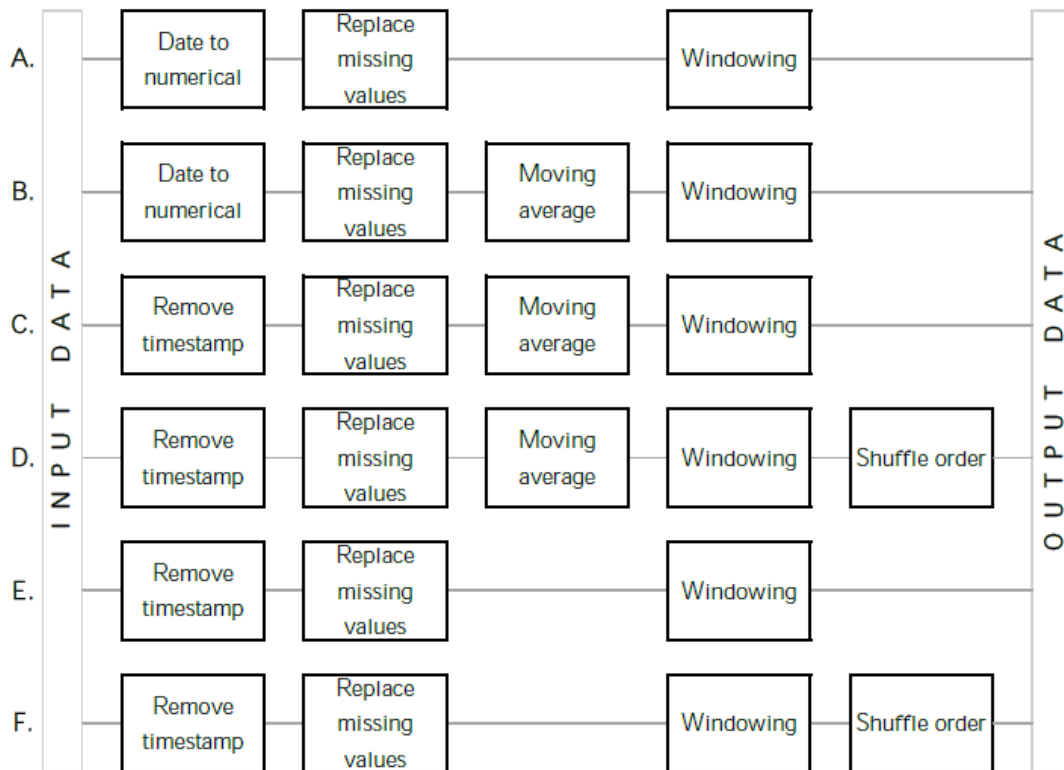
در این مقاله از نرم افزار RapidMiner version 7.6 (RapidMiner, 2017) استفاده شد.

بوسیله این نرم افزار میتوان به آماده سازی داده‌ها، یادگیری ماشین، یادگیری عمیق، استخراج متن و تحلیل پیش بینی پرداخت.

Deep Learning در این تحقیق نیز بوسیله این نرم افزار به کارگیری شد

- حالت های مختلف پیش پردازش

در این مطالعه داده ها را در ۶ حالت مختلف پیش پردازش کرده و از آنها جهت فرآیند تخمین با هوش مصنوعی استفاده شد تا تاثیر پیش پردازش داده ها بر روی نتایج تخمین مشخص گردد.



شکل ۲-۱۲: شش حالت مختلف برای پیش پردازش داده ها در این مقاله

- محاسبه‌ی خطای محاسبات :

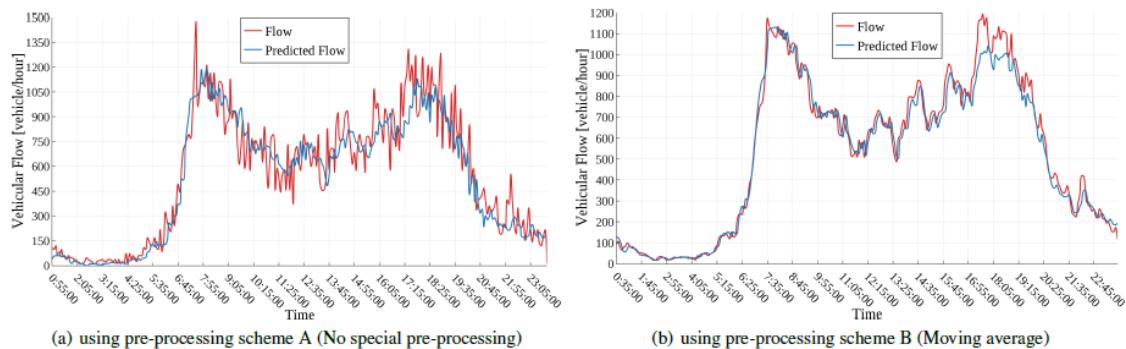
جهت ارزیابی نتایج تخمین از دو روش خطای جذر میانگین مربعات Root Mean Square Error

(RMSE) و میانگین خطای نسبی (MRE) استفاده شد.

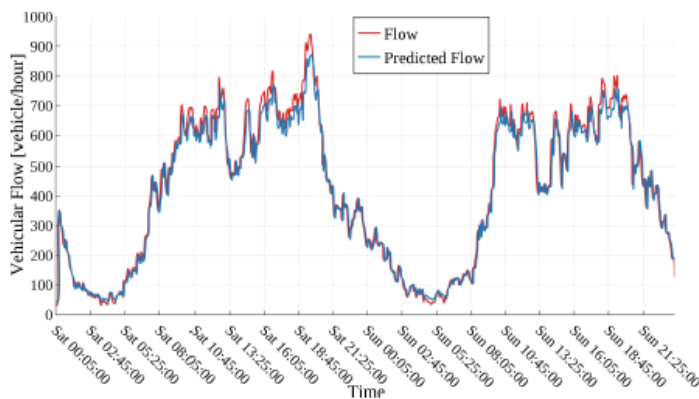
جدول ۱۰-۱: خلاصه‌ی خطای تخمین جریان وسایل نقلیه بوسیله یادگیری عمیق در داده های مختلف با پیش پردازش های متفاوت

Training data	Testing data	Moving average	No timestamp input	Shuffled order	Window size [min]	RMSE	MRE	Correlation
2 weekdays	1 weekday				55	134.50	24.50%	93.60%
		✓			45	47.67	7.36%	99.33%
		✓	✓		60	40.41	11.00%	99.53%
		✓	✓	✓	15	45.73	8.92%	99.24%
			✓		20	126.51	28.52%	94.31%
			✓	✓	60	135.29	28.85%	94.27%
2 weekends	1 weekend				30	122.93	27.22%	91.81%
		✓			30	26.69	9.05%	99.42%
		✓	✓		25	24.22	5.39%	99.53%
		✓	✓	✓	35	35.01	8.23%	98.99%
			✓		15	100.11	27.27%	92.60%
			✓	✓	20	97.66	27.75%	92.71%
2 weeks	1 week				40	129.99	29.03%	93.54%
		✓			40	21.05	5.80%	99.84%
		✓	✓		25	43.50	8.79%	99.27%
		✓	✓	✓	30	42.34	8.82%	99.27%
			✓		35	118.40	27.60%	94.56%
			✓	✓	60	121.98	28.47%	94.23%

از میان پیش پردازش های انجام شده، میانگین متحرک بهبود قابل توجهی در عملکرد مدل داشت و مقادیر RMSE, MRE and Correlation در هنگام استفاده از میانگین متحرک 36.29, 8.15% و 99.38% و بدون استفاده از میانگین متحرک 128/02, 27.69% , 93.51% بودند (شکل ۲-۱۳).



شکل ۲-۱۳: جریان واقعی و جریان تخمین زده شده در یک روز هفته (a) با پیش پردازش میانگین متحرک و (b) بدون پیش پردازش میانگین متحرک



شکل ۲-۱۴: جریان واقعی و جریان تخمین زده شده در یک روز آخر هفته (شنبه یکشنبه) با اعمال پیش پردازش میانگین متحرک بر داده ها

- ✓ از میان انواع پیش پردازش ها، میانگین متحرک موثرترین پیش پردازش است
- ✓ نتایج نشان می دهد که Deep Learning تنها با دو لایه پنهان از ۵۰ واحد که هر یک با داده های ترافیکی یک هفته آموزش دیده اند می تواند جریان را در هر روز از هفته و همچنین تعطیلات با دقت قابل توجه پیش بینی کند

۲-۹ نتیجه گیری

- با توجه به مقاله Hichem Omrani می توان نتیجه گرفت که در تخمین داده های مشابه تحقیق انجام شده در این پایان نامه artificial neural net-MLP در مقایسه با روش های artificial neural net-RBF, multinomial logistic regression, and support vector machines عملکرد مناسبی خواهد داشت.
- با توجه به مقاله Gusri Yaldi جهت آموزش شبکه عصبی، الگوریتم Levenberg-Marquardt (LM) در مقایسه با الگوریتم های Back propagation (BP) و Variable Learning Rate (VLR) از عملکرد بهتری در تخمین ماتریس سفر برخوردار است.
- با توجه به مقاله Albertengo پیش از استفاده از داده های خام جمع آوری شده در

فرآیند تعلیم هوش مصنوعی و تخمین پارامترهای مورد نظر، می‌بایست داده‌های زمانی و

مکانی را با استفاده از روش‌های زیر پیش پردازش نمود تا نتایج مناسبی حاصل گردد:

Date to numerical, Replace missing values, Time series windowing, Shuffled order

فصل ۳ : مبانی هوش مصنوعی

۳-۱ مقدمه

هنوز تعریف دقیقی که مورد قبول همه دانشمندان این علم باشد برای هوش مصنوعی ارائه نشده است. اما شاید بتوان دقیق ترین تعریف را این دانست که: هوش مصنوعی عبارت است از مطالعه این که چگونه کامپیوترها را می توان وادار به کارهایی کرد که در حال حاضر انسان ها آنها را بهتر انجام می دهند.

در این فصل به بررسی مبانی هوش مصنوعی و خصوصا زیرشاخه های اصلی آن یعنی شبکه عصبی مصنوعی (Artificial Neural Networks) و منطق فازی (Fuzzy Logic) پرداخته شده است.

۳-۲ شبکه عصبی مصنوعی

شبکه های عصبی را می توان با اغماض زیاد، مدل های الکترونیکی از ساختار عصبی مغز انسان نامید. مکانیسم فراگیری و آموزش مغز اساسا بر تجربه استوار است و مدل های الکترونیکی شبکه های عصبی طبیعی نیز بر اساس همین الگو بنا شده اند و روش برخورد چنین مدل هایی با مسائل، با روش های محاسباتی که به طور معمول توسط سیستم های کامپیوتری در پیش گرفته شده اند، تفاوت دارد. شبکه های عصبی شبیه سازی شده یا کامپیوتری، فقط قادرند تا بخش کوچکی از خصوصیات و ویژگی های شبکه های عصبی بیولوژیک را شبیه سازی کنند. در حقیقت، هدف از ایجاد یک شبکه عصبی نرم افزاری، بیش از آنکه شبیه سازی مغز انسان باشد، ایجاد مکانیسمی برای حل مسائل مهندسی با الهام از الگوی رفتاری شبکه های بیولوژیک است. در شبکه های عصبی بیولوژیک، نرون ها در ساختاری سه بعدی به یکدیگر اتصال یافته اند. اتصالات بین نرون ها در شبکه های عصبی بیولوژیک آنقدر زیاد و پیچیده است که به هیچ وجه نمی توان شبکه مصنوعی مشابه آن طراحی کرد. تکنولوژی مدارات مجتمع امروزی به ما امکان می دهد که شبکه های عصبی را در ساختار های دو بعدی طراحی کنیم.

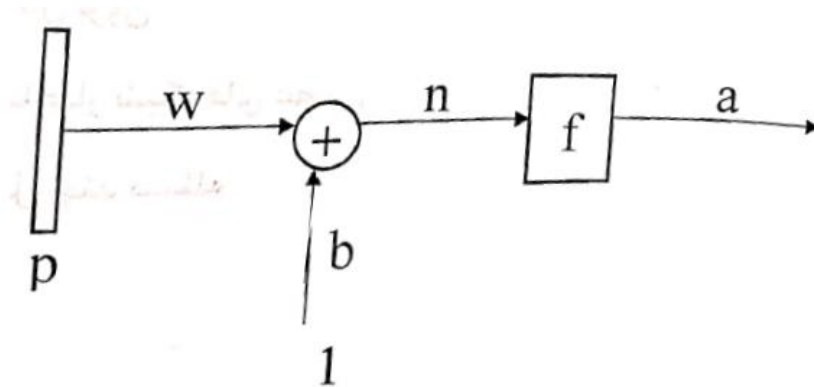
۱-۲-۳ مدل ریاضی نرون

نرون (شکل ۱-۳) کوچکترین واحد پردازشگر اطلاعات است که اساس عملکرد شبکه های عصبی را

تشکیل می دهد.

شکل زیر ساختار یک نرون تک ورودی را نشان می دهد. اسکالرهایی p و a به ترتیب ورودی و

خروجی می باشند.



شکل ۱-۳ ساختار یک نرون تک ورودی

میزان تاثیر p روی a بوسیله اسکالر w (وزن) تعیین می شود. ورودی دیگر که مقدار ثابت ۱ است،

در جمله بایاس b ضرب شده و سپس با wp جمع می شود، این حاصل جمع، ورودی ناخالص n برای

تابع محرک (یا تابع تبدیل) f خواهد بود. بدین ترتیب خروجی نرون با معادله زیر تعریف می شود:

$$a=f(wp+b)$$

باید توجه داشت که پارامترهای w و b قابل تنظیم هستند و تابع محرک f نیز توسط طراح انتخاب

می شود. بر اساس انتخاب f و نوع الگوریتم یادگیری، پارامترهای w و b تنظیم می شوند. یادگیری

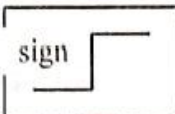
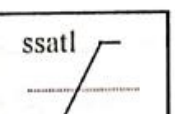
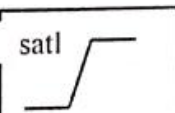
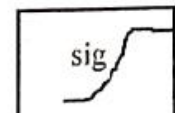
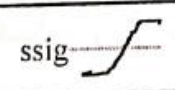
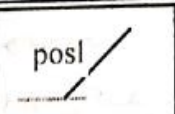
بدین معنی است که w و b طوری تغییر می کنند که رابطه ورودی و خروجی نرون با هدف خاصی

مطابقت نماید.

۲-۲-۳ توابع محرک

تابع محرک f می تواند خطی یا غیر خطی باشد. یک تابع محرک بر اساس نیاز خاص حل یک مسئله - مسئله ای که قرار است به وسیله شبکه عصبی حل شود - انتخاب می شود. در عمل تعداد محدودی از توابع محرک مورد استفاده قرار می گیرند، که در جدول ۱-۳ لیست شده است.

جدول ۱-۳ توابع محرک با علائم قراردادی

ردیف	نام	تعریف تابع	علائم قراردادی
۱	آستانه‌ای دو مقداره	$a = 0, n < 0$ $a = 1, n \geq 0$	
۲	آستانه‌ای دو مقداره متقارن	$a = -1, n < 0$ $a = 1, n \geq 0$	
۳	خطی	$a = n$	
۴	آستانه‌ای خطی متقارن	$a = -1, n < -1$ $a = n, -1 \leq n \leq 1$ $a = 1, n > 1$	
۵	آستانه‌ای خطی	$a = 0, n < 0$ $a = n, 0 \leq n \leq 1$ $a = 1, n > 1$	
۶	زیگموئیدی	$a = \frac{1}{1 + e^{-n}}$	
۷	تانژانت هیپربولیکی	$a = \frac{e^n - e^{-n}}{e^n + e^{-n}}$	
۸	خطی مثبت	$a = 0, n < 0$ $a = n, n \geq 0$	

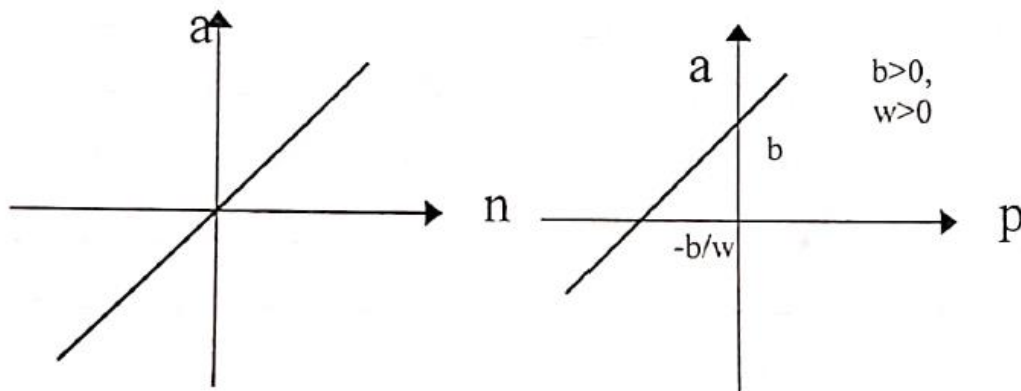
در این جا به چند مورد از مهمترین آنها اشاره می کنیم

۱- تابع محرک خطی

خروجی این تابع برابر ورودی آن است:

$$a = f_1(n) = n$$

تابع خطی در شکل (۲-۳) نمایش داده شده است. از نرونهايي با تابع تبدیل فوق، در شبکه های خاصی مانند آدالاین استفاده شده است. این شبکه ها در فصل هفتم بطور کامل مورد بررسی قرار خواهند گرفت. اهمیت جمله بایاس ۶ را در شکل (۲-۲) می بینیم. اگر پاسخ نرون، a ، بر حسب ورودی ۲ رسم شده باشد، جمله بایاس b موجب جابه جایی منحنی در فضای ورودی می گردد و به عبارتی موجب می گردد که نرون به زیر فضایی از فضای ورودی بایاس گردد، که خود انتخاب کلمه بایاس را برای ترم b توجیه می کند.

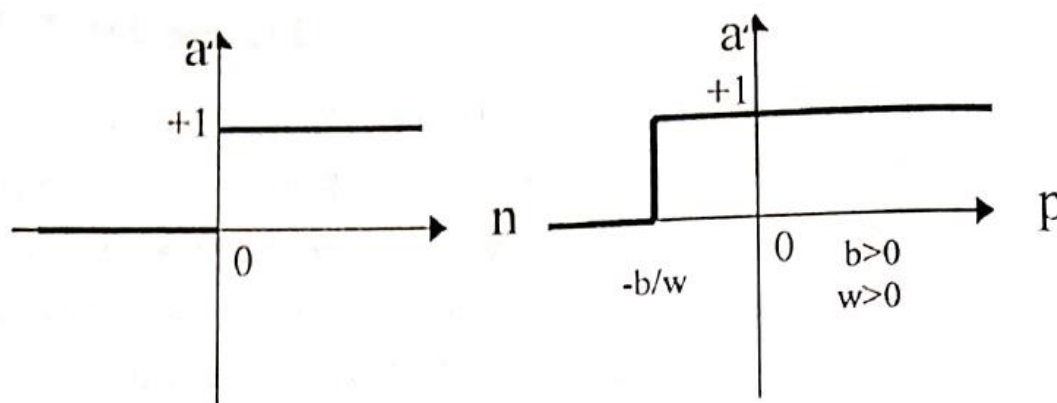


شکل ۲-۳ تابع محرک خطی

۲- تابع محرک آستانه ای دو مقداره حدی

این تابع در شکل (۳-۳) نشان داده شده است. همان گونه که مشاهده میشود، مقدار خروجی ۰ یا ۱ است. اگر آرگومان n کوچکتر از ۰ باشد و یا به عبارتی ورودی p کوچکتر از باشد، مقدار تابع ۰ است و در غیر این صورت خروجی نرون برابر ۱ خواهد شد. عموماً تابع محرک، دامنه خروجی نرون را

محدود می سازد و به همین علت آن را تابع محدودساز^۱ نیز می نامند. خروجی نرون معمولاً برای این گونه توابع ، در بازه متناهی $[10, 1]$ یا $[-1, 1]$ قرار دارد، که در حالت اخیر تابع را تابع محرک آستانهای دو مقداره حدی گویند.



شکل ۳-۳ تابع محرک آستانهای دو مقداره حدی

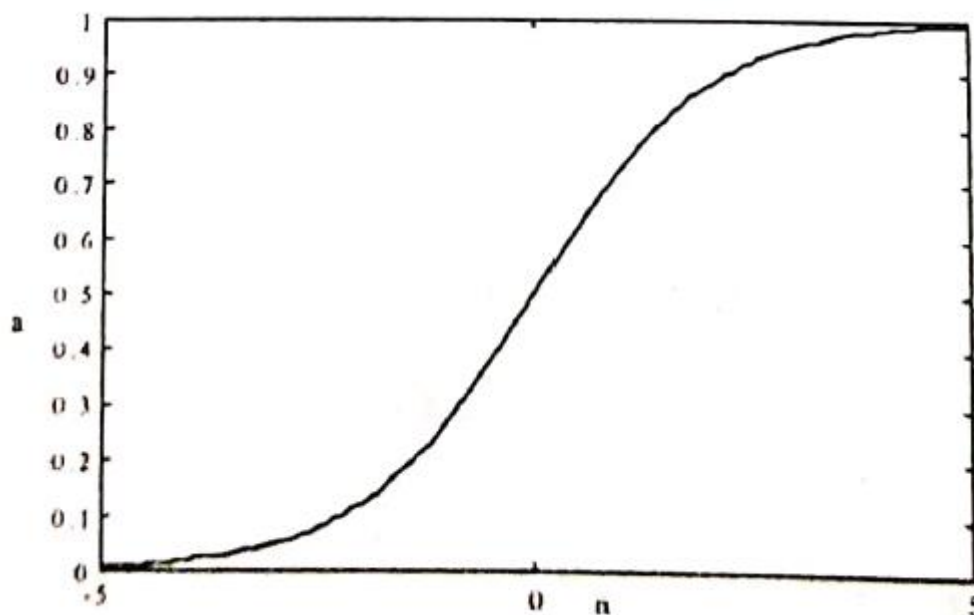
۳- تابع محرک زیگموئید

این تابع با فرمول کلی زیر بیان می شود:

$$a = f_s(n) = \frac{1}{1 + e^{-cn}} \quad , \quad c > 0$$

شکل این تابع به ازاء $C = 1$ در تصویر (۴-۳) رسم شده است. مقدار C وسعت ناحیه خطی بودن تابع را تعیین می کند. مثلاً اگر C خیلی بزرگ باشد، شکل منحنی به تابع محرک آستانهای دو مقداره حدی نزدیکتر میشود. این تابع در شبکه های عصبی مورد استفاده زیادی دارد که به عنوان مثال، می توان به شبکه های عصبی چند لایه با قانون یادگیری پس انتشار خطا اشاره کرد.

¹ Hardlimit or squash



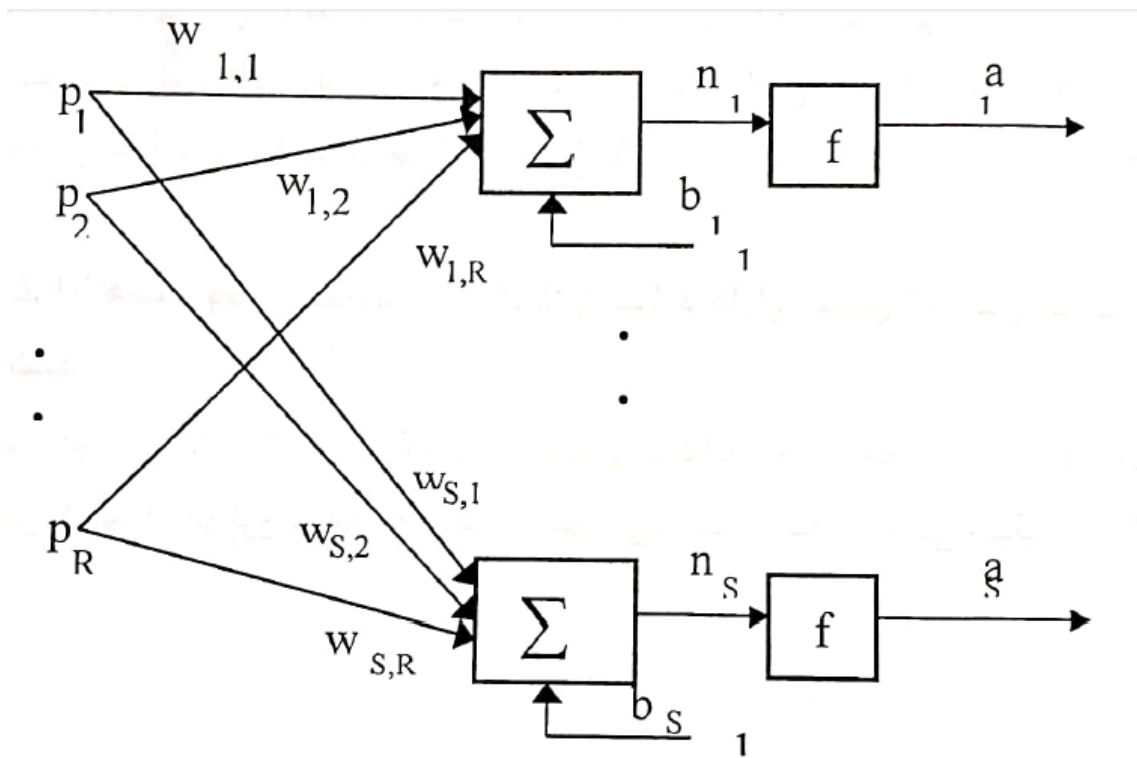
شکل ۳-۴ تابع محرک زیگموئید

۳-۲-۳ ساختار شبکه های عصبی

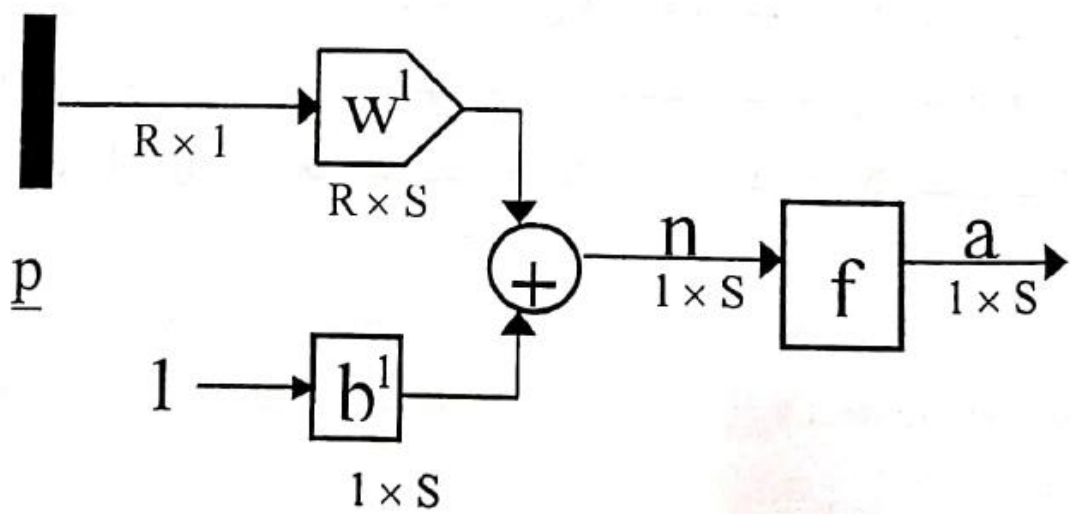
باید توجه داشت که معمولا نرونی با ورودی های زیاد نیز، به تنهایی برای حل مساله فنی - مهندسی کفایت نمی کند. مثلا برای مدلسازی نگاشتهایی که دارای دو خروجی هستند احتیاج به دو نرون داریم، که بطور موازی عمل کنند. در این حالت یک لایه خواهیم داشت که از اجتماع چند نرون تشکیل شده است.

• شبکه تک لایه

یک شبکه تک لایه با S نرون در شکل (۳۲-۵) نشان داده شده است. شکل مذکور را می توان به فرم فشرده تصویر (۳-۶) هم نمایش داد. ورودی شبکه با بردار p و خروجی آن با بردار a نشان داده شده است. باید توجه داشت که هر یک از ورودیها به همه نرونها متصل شده است. ماتریس W نیز در این حالت دارای S سطر و R ستون می باشد. همان گونه که در شکل مشاهده می شود، لایه ها شامل ماتریس وزن، جمع کننده ها، بردار بایاس b (دارای S عنصر) و تابع تبدیل f هستند.



شکل ۳-۵ شبکه تک لایه با S نرون



شکل ۳-۶ فرم فشرده یا ماتریسی شبکه تک لایه با S نرون

۳-۲-۱ شبکه های چند لایه

در شبکه های تک لایه، بردار ورودی (p) توسط نرون های لایه (عناصر محاسباتی)، طبق رابطه

$$a=f(Wp+b)$$

به بردار خروجی مرتبط می شود. این شبکه شکل ساده ای از شبکه های پیشخور^۱ می باشد. هر

لایه ماتریس وزن W ، بردار بایاس b ، بردار ورودی خالص n و بردار خروجی مختص خود را دارد. جهت

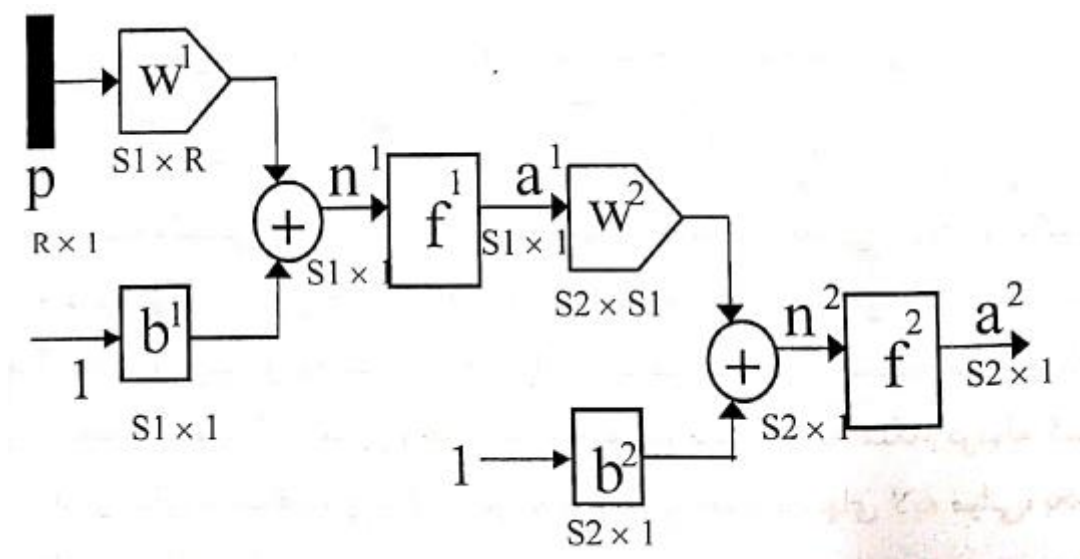
تمایز متغیرهای فوق و این که کدام متغیر به کدام لایه تعلق دارد، نیاز داریم که علامت دیگری را هم

به متغیرهای فوق تخصیص دهیم. از این رو از بالا مشخص نمودن لایه استفاده می کنیم، بنابراین

ماتریس وزن را برای لایه اول با W^1 مشخص می نماییم، یک شبکه پیشخور دو لایه را می توان به

شکل (۳-۸) ترسیم نمود. همان گونه که از شکل پیداست، تعداد R ورودی و تعداد S^1 نرون در لایه

اول و تعداد S^2 نرون در لایه دوم در شبکه چند لایه پیشخور موجود است.

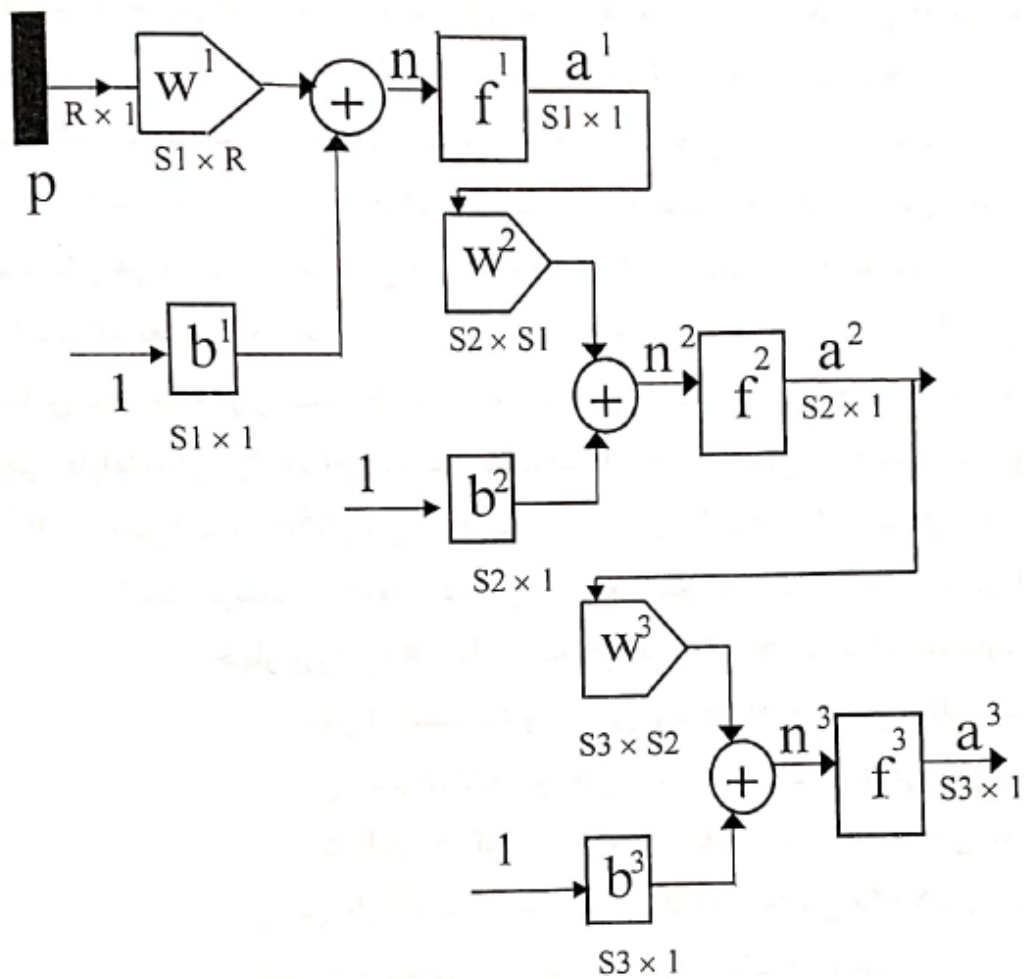


شکل ۳-۷ شبکه پیشخور دو لایه

در حالات پیشرفته تر می توان تعداد لایه های میانی را افزایش داد و برای مثال در شکل ۳-۹ تصویر

یک شبکه عصبی پیشخور با دولایه میانی مشاهده می شود.

^۱ feedforward



شکل ۳-۸ فرم ساده شده شبکه پیش‌خور سه لایه

۳-۲-۴ لایه خروجی

لایه ای که خروجی آن، خروجی نهایی شبکه باشد، به لایه خروجی موسوم است. لایه های دیگر به لایه های میانی موسومند. شبکه موجود در شکل (۲-۱۰) دارای یک لایه میانی است. شکل (۲-۱۱) فرم ساده شده یک شبکه عصبی پیش‌خور را با دو لایه میانی نمایش می دهد. در این جا لایه خروجی با ماتریس وزن W^3 ، بردار بایاس b^3 و تابع محرک f^3 مشخص می شود. لایه میانی اول با ماتریس وزن W^1 ، بردار بایاس b^1 و تابع محرک f^1 و لایه میانی دوم با ماتریس وزن W^2 ، بردار بایاس b^2 و تابع محرک f^2 مشخص می شوند.

۳-۳ روش های تخمین مبتنی بر هوش مصنوعی

۳-۳-۱ مقدمه

هوش مصنوعی از روش های گوناگونی جهت تخمین داده ها استفاده می کند که از پرکاربردترین آنها می توان به روش شبکه عصبی^۱ و روش های فازی^۲ اشاره کرد.

۳-۳-۲ شبکه عصبی

در این روش، ابتدا داده هایی بعنوان داده های ورودی به شبکه عصبی داده می شود و با استفاده از روش انتخابی ما، داده ها را به سه دسته ی داده های آموزش^۳، داده های ارزیابی^۴ و داده های تست^۵ تقسیم می کند. سپس با استفاده از تنظیمات انجام شده از قبل الگوریتم آموزش تعیین شده و تعداد لایه های مشخص شده و ویژگی های دیگر، ابتدا بر اساس داده های آموزش سعی می کند تا رابطه ای میان داده های ورودی بیابد، سپس این رابطه را بر اساس تعداد دفعات تعیین شده، با استفاده از داده های ارزیابی آزمایش می کند تا میزان دقت تخمین مشخص گردد که آیا رابطه بدست آمده از دقت کافی برخوردار است یا خیر؟ و در نهایت، بهترین نتیجه را با استفاده از داده های تست آزمایش می کند تا دقت بدست آمده نیز مورد ارزیابی مجدد قرار گیرد.

در ذیل نکاتی راجع به این فرآیند از جمله استراتژی های آموزش شبکه عصبی، انواع شبکه های عصبی، ویژگی ها و کاربردها بررسی می شود.

استراتژی مختلف برای آموزش شبکه عصبی عبارتند از:

^۱ Artificial neural network

^۲ Fuzzy

^۳ Training data

^۴ Validation data

^۵ Test data

۱-۲-۳-۳ یادگیری بدون نظارت^۱

در روش یادگیری بدون نظارت داده های اولیه یا ورودی های شبکه بدون پاسخ یا هدف (Target) هستند و وزن ها و بایاس شبکه فقط بر اساس ورودی ها اصلاح می گردند. بیشتر روش های یادگیری بدون نظارت برای خوشه بندی^۲ استفاده می شوند که در آن یافتن ساختار و رابطه درون مجموعه داده ها بررسی می شود.

۱-۲-۳-۳ یادگیری با نظارت^۳

یک مجموعه داده به عنوان مجموعه آموزش^۴ با پاسخ صحیح برای آموزش شبکه استفاده می شود. قانون یادگیری^۵ بر اساس اصلاح و تنظیم وزن ها و بایاس^۶ شبکه برای نزدیکی به مقادیر پاسخ تولید شده و از این قانون برای پیش بینی داده های جدید استفاده می شود. روش یادگیری با نظارت برای مدل سازی، طبقه بندی، رگرسیون و درون یابی به کار می رود.

^۱ unsupervised learning

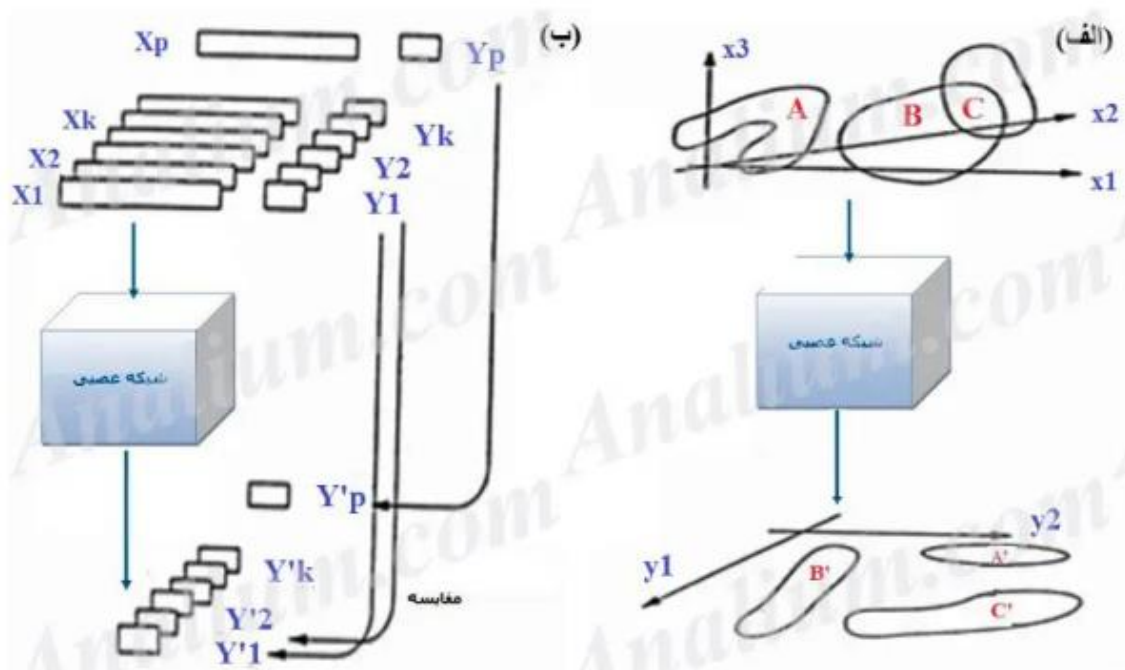
^۲ Clustering

^۳ supervised learning

^۴ training set

^۵ learning rule

^۶ bias



شکل ۳-۹: (الف) یادگیری بدون نظارت و (ب) یادگیری با نظارت

۳-۲-۳ یادگیری تقویتی^۱

در این شبکه ها آموزش از طریق مشاهده و سعی و خطا صورت می گیرد. در واقع با توجه به تعاملی که با محیط پیرامون خود دارد عملی انجام داده و نتیجه عمل را به صورت پاداش و تنبیه برای نزدیکی و دوری به هدف مشخص می شود. شبکه کارهایی انجام می دهد که منجر به کسب پاداش بیشتری می شود.

۳-۳-۳ انواع شبکه عصبی

انواع مختلفی از شبکه های عصبی تاکنون ارائه شده و در زمینه های مختلفی به کار رفته و عملکرد موفقی داشته اند. در ادامه چند نوع متداول و با اهمیت شبکه عصبی معرفی می شود.

¹ Reinforcement learning

۳-۳-۱ شبکه عصبی پیش خور^۱

در این نوع شبکه پردازش داده های ورودی رو به جلو انجام می شود و مسیر پردازش به نوروں های لایه قبلی بر نمی گردد و خروجی هر لایه فقط بر لایه بعدی تاثیرگذار خواهد بود. یعنی سیگنال یک طرفه خواهد بود. در این نوع شبکه مقدار پارامتر خروجی، براساس پارامترهای ورودی و یک سری وزن های اولیه تعیین می گردد. مقادیر ورودی با هم ترکیب می شوند و سپس در لایه های پنهان^۲ از آنها استفاده می شوند و در آخر مقادیر لایه های پنهان نیز برای محاسبه مقادیر خروجی ترکیب می شوند.

۳-۳-۲ شبکه عصبی انتشار برگشتی^۳

در این نوع شبکه، خطای خروجی با مقایسه مقدار خروجی با مقدار تجربی یا مقدار مورد نظر محاسبه می گردد. از مقدار و میزان خطا برای تصحیح شبکه و تغییر وزن ها استفاده می شود. این محاسبات از گره خروجی شروع شده و به لایه های پنهان ادامه می یابد. این عمل به ازای تمامی داده های اولیه تکرار می گردد. به هربار اجرای این الگوریتم برای تمام داده ها یک دوره^۴ گفته می شود. دوره آنقدر تکرار می شوند تا مقدار خطا به مقدار ثابتی رسیده و تغییر نکند.

۳-۳-۴ ویژگی های روش های شبکه عصبی

- پردازش موازی و در نتیجه سرعت بالای محاسبات و بدون نویز^۵ دارند.
- توانایی تعمیم پذیری و تطبیق پذیری
- پردازش خطی و غیر خطی

^۱ Feed-Forward

^۲ hidden layer

^۳ Backpropagation

^۴ epoch

^۵ noise

- دامنه کاربردی وسیع شامل پردازش سیگنال، دسته بندی، شناسایی الگو، رگرسیون، مدل سازی و ..
- بیشتر شبکه های عصبی مشابه جعبه سیاه^۱ هستند یعنی روابط بین ورودی و خروجی و پروسه عملکرد شبکه برای کاربر ناشناخته و غیرقابل تفسیر است.
- زمان محاسبه نسبتاً طولانی دارند مخصوصاً برای حل مسائل پیچیده با شبکه های چند لایه. اگرچه با پیشرفت قدرت محاسبه کامپیوترهای امروزی و استفاده از سیستم های موازی^۲ این مشکل چندان جدی گرفته نمی شود.

۵-۳-۳ زمینه های کاربردی روش های شبکه عصبی

- تخمین داده های آماری
- کنترل، بهینه سازی و مدل سازی فرایندهای صنعتی
- تشخیص و ارزیابی پدیده های پزشکی، پیش بینی هزینه درمان و اقامت بیمار و ...
- مدل سازی، پیش بینی، دسته بندی آنالیز سری های زمانی و ارزیابی داده ها
- الگو شناسی، ارزیابی اکوسیستم، شناسایی ژن ها، تجزیه و تحلیل سیستم های بیولوژیک، شناسایی ترکیبات شیمیایی و ...
- مدیریت ریسک ، بازاریابی هدف و تحقیق مشتری

۶-۳-۳ نرم افزارهای متداول برای محاسبات ANN

- nntool Matlab toolbox
- neurosolutions
- Simbrain

^۱ Black boxes

^۲ parallel systems

فصل ۴ : روش تحقیق

۴-۱ مقدمه

در این تحقیق هدف آن است که بوسیله شبکه عصبی مصنوعی و با استفاده از نرم افزار MATLAB R2013a با توجه به طبقه بندی های مختلف توزیع جمعیتی در سطح شهر و همچنین با توجه به سطح زیربنای کاربری اراضی موجود در هر ناحیه شهری تخمین دقیقی از ماتریس سفر بین نواحی شهر بدست آورد. بدین منظور ابتدا داده های مورد نیاز از سازمان ترافیک و سازمان فناوری اطلاعات و ارتباطات شهرداری شهرهای مشهد و اصفهان جمع آوری شد و پس از آن بطور جداگانه برای هر کدام از شهرها، داده ها بوسیله نرم افزار SPSS پردازش شده و ضریب همبستگی شان استخراج شد. سپس ضرایب همبستگی میان پارامترهای جمع آوری شده، بوسیله نرم افزار SPSS و نرم افزار MATLAB محاسبه شد و در هر کدام از مقادیر ضریب همبستگی، میان پارامترهای همبسته، پارامتری که دامنه تغییر مقادیر داده های بالاتری داشت (واریانس داده های بالاتر) در تحقیق مورد استفاده قرار گرفت. در این فصل روش تحقیق توضیح داده خواهد شد و نتایج تخمین در فصل ۵ طرح و بررسی خواهند شد.

۴-۲ جمع آوری داده ها

۴-۲-۱ جمع آوری داده های شهر مشهد

(۱) ماتریس سفر سال ۱۳۹۵

بر اساس تقسیم بندی نواحی ترافیکی (۲۵۳ ناحیه) شهر مشهد در سال ۱۳۹۵، ماتریس سفر از طریق سازمان ترافیک شهرداری مشهد بدست آمد، این ماتریس سفر بصورت فایل excel ای حاوی ۲۵۳ ردیف و ۲۵۳ ستون است و در CD ضمیمه این تحقیق با نام ((ماتریس سفر شهر مشهد سال ۱۳۹۵ بر اساس نواحی ترافیکی)) آورده شده است.

۲) داده های جمعیتی سال ۱۳۹۵

بر اساس سرشماری جمعیتی انجام شده در سال ۱۳۹۵ شهر مشهد، داده های مربوط به سرشماری جمعیتی شهر مشهد از طریق اداره فناوری اطلاعات و ارتباطات شهرداری مشهد بدست آمد، داده های سرشماری بصورت فایل excel ای حاوی ۲۰۱۰۸ ردیف و ۱۵۲ ستون است و در CD ضمیمه این تحقیق با نام ((سرشماری جمعیتی مشهد سال ۱۳۹۵)) آورده شده است.

۳) داده های آماری کاربری اراضی سال ۱۳۹۵ میان نواحی مختلف شهر مشهد بر اساس

تقسیم بندی نواحی شهرداری

بر اساس داده های مربوط به اداره فناوری اطلاعات و ارتباطات شهرداری مشهد، کاربری اراضی موجود در هر ناحیه شهرداری (۱۳ گانه) موجود است، داده های سطح زیربنای کاربری اراضی شهر مشهد بصورت فایل excel ای حاوی ۳۷۰۶۴ ردیف و ۶ ستون است و در CD ضمیمه این تحقیق با نام ((کاربری اراضی مشهد داده های اولیه)) آورده شده است.

۴-۲-۲ جمع آوری داده های شهر اصفهان

۱) ماتریس سفر سال ۱۳۹۵

بر اساس تقسیم بندی نواحی ۱۵ گانه ترافیکی شهر اصفهان، ماتریس سفر سال ۱۳۹۵ این شهر بصورت فایل اکسلی حاوی ۱۶ ردیف و ۱۶ ستون موجود است که از طریق سازمان فناوری اطلاعات و ارتباطات شهرداری اصفهان تهیه شد. این ماتریس در جدول ۴-۱ آورده شده است:

جدول ۴-۱ بخشی از ماتریس سفر سال ۱۳۹۵ اصفهان بر اساس نواحی شهرداری

تعداد سفر	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰
۱	۸۰۰۴۳	۲۶۴۵۷	۳۴۱۲۰	۹۲۰۵	۳۱۸۸۴	۱۵۵۷۹	۲۴۹۵۵	۶۰۰۸۳	۲۷۹۷۲	۱۲۲۸۶
۲	۲۷۶۴۹	۱۰۸۷۴۰	۱۰۲۹۸	۲۵۲۶	۱۰۰۵۸	۵۹۹۲	۷۳۱۶	۵۴۱۵۴	۱۳۹۷۲	۲۷۴۵
۳	۳۳۳۰۹	۹۹۹۴	۱۲۸۳۹۱	۵۷۸۱۰	۳۲۵۰۹	۳۶۲۸۳	۶۶۴۵۶	۴۶۸۱۷	۷۷۱۵	۶۵۰۴۸
۴	۹۴۱۲	۲۵۷۰	۵۹۷۷۶	۱۰۶۸۸۸	۱۸۳۶۹	۱۹۷۵۲	۱۱۴۵۸	۱۰۱۸۹	۱۵۰۴	۲۲۹۴۶
۵	۳۳۵۷۵	۹۵۳۴	۳۵۲۸۴	۱۸۴۹۰	۴۴۷۰۷۹	۹۹۴۴۱	۲۱۶۳۲	۳۳۷۹۹	۹۴۳۲	۱۳۵۲۱
۶	۱۶۲۱۱	۶۰۷۱	۳۷۰۱۹	۱۸۹۱۷	۹۹۴۷۱	۱۴۰۸۴۵	۱۸۴۷۵	۲۰۷۳۹	۴۳۵۱	۱۴۵۱۱
۷	۲۶۰۲۴	۷۳۹۰	۶۹۹۶۹	۱۱۶۹۵	۲۳۴۰۱	۱۸۴۶۸	۲۳۷۹۵۴	۶۶۲۰۶	۲۳۱۲	۵۸۲۰۴
۸	۶۲۷۳۴	۵۴۵۴۵	۵۰۴۴۰	۱۰۸۸۹	۳۵۳۲۱	۲۱۴۷۹	۶۵۷۷۸	۴۰۹۵۱۳	۸۵۱۹	۱۹۳۷۸
۹	۲۸۵۵۳	۱۴۱۴۰	۷۵۶۹	۱۵۶۴	۹۱۵۰	۴۰۹۹	۲۱۲۲	۸۱۳۰	۴۷۵۶۶	۱۶۹۴
۱۰	۱۲۴۰۵	۲۶۵۹	۶۸۷۵۵	۲۲۹۸۳	۱۳۱۱۸	۱۵۱۲۴	۵۵۸۲۴	۱۸۹۱۲	۱۶۴۱	۱۵۱۵۷۴

۲) داده های آماری سرشماری سال ۱۳۹۵

داده های مربوط به جمعیت ساکن در هر ناحیه شهرداری اصفهان از طریق سازمان فناوری اطلاعات و ارتباطات شهرداری اصفهان تهیه شد که بصورت فایل excel حاوی ۱۶ ردیف و ۲۰ ستون می باشد که در CD ضمیمه این تحقیق با نام ((جمعیت اصفهان بر اساس سرشماری سال ۹۵)) آورده شده است. بخشی از این فایل در جدول ۴-۵ قابل مشاهده است:

جدول ۴-۲ بخشی از فایل سرشماری اصفهان در سال ۱۳۹۵ بر اساس ۱۵ ناحیه شهرداری

مناطق شهرداری	جمعیت کل	جمعیت مرد	جمعیت زن	تعداد کل خانوار	جمعیت ۶ ساله و بیشتر	جمعیت ۶ ساله و بیشتر یاسواد	جمعیت ۶ ساله و بیشتری سواد	مرد ۶ ساله و بیشتر	جمعیت ۱۰ ساله و بیشتر	جمعیت ۱۰ ساله و بیشتر مرد
1	79091	38589	40214	27187	73891	70023	3860	36076	70761	34420
2	69120	34663	33458	21497	61430	54992	6429	31185	57412	29156
3	110368	54008	55609	36791	102271	95587	6623	50293	97469	47839
4	133076.499	65917	68041	43277	122991	115826	7155	61210	117060	58198
5	150865	72892	77646	47934	142824	138254	4507	69043	137054	65987
6	112042.7978	54776	57004	37509	105400	100337	5052	51529	100814	49200
7	168292.5025	84761	83341	52738	153240	142702	10535	77084	144052	72369
8	239756	120091	119101	75157	219504	204239	15247	109890	206692	103301
9	72252.57021	36186	35362	23217	65443	59244	6197	33109	61815	31220
10	207687.9696	104743	102890	65571	188748	171876	16862	95090	177091	89107
11	58804.9847	29955	28825	17939	52967	46166	6799	26980	49358	25143
12	134388.085	67270	66245	41898	121271	113783	7477	61007	113600	57049
13	127793.5056	69289	58245	37283	118786	113212	5270	64804	113279	61964
14	160761.2332	81760	78342	46676	142265	123900	18360	72663	130906	66855
15	120369.4256	60732	57552	36812	106838	93777	13047	54911	99870	51389

۳) داده های آماری کاربری اراضی سال ۱۳۹۵ میان نواحی مختلف شهر اصفهان بر اساس

تقسیم بندی نواحی شهرداری

داده های مربوط به سطح زیربنای کاربری اراضی مختلف واقع در هریک از نواحی شهرداری اصفهان، از طریق سازمان فناوری اطلاعات و ارتباطات شهرداری اصفهان بدست آمد. این داده ها

مربوط به نقشه های تهیه شده در سال ۱۳۹۵ بود و در فایل JPEG آمده است. در این فایل، عدد متناظر برای هر کاربری بصورت سهم (درصد) آن کاربری از مجموع سطح آن ناحیه بیان شده است. این فایل در جدول ۴-۶ نمایش داده شده است و در فایلی بصورت جداگانه نیز سطح هر ناحیه شهرداری به همراه سطح تعدادی از کاربری های واقع در آن ناحیه آورده شده است که نمونه ای از این فایل ها که در فرمت JPEG هستند نیز در جدول ۴-۷ آورده شده است.

جدول ۴-۳ کاربری اراضی موجود در نواحی مختلف شهرداری اصفهان بر اساس درصد

سهم (درصد)											کاربری
منطقه ۱	منطقه ۲ و ۱۱	منطقه ۳	منطقه ۴	منطقه ۵ و ۱۳	منطقه ۶	منطقه ۷ و ۱۴	منطقه ۸ و ۱۲	منطقه ۹	منطقه ۱۰	منطقه ۱۵	
۴۸,۸۴	۲۷,۰۹	۴۴,۹۸	۳۸,۴۳	۲۶,۵۲	۲۵,۵۸	۲۸,۶۵	۲۷,۸۴	۴۵,۵۷	۳۸,۹۸	۴۳,۸	مسکونی
۳,۸۳	۱,۶۸	۵,۲۸	۱,۰۰	۰,۸۷	۱,۱۸	۱,۰۶	۱,۰۱	۱,۱۹	۱,۳۱	۱,۴	مختلط تجاری
۱,۳۹	۰,۷۵	۱,۹۴	۱,۳۰	۰,۸۱	۱,۱۶	۰,۸۳	۰,۷۸	۰,۸۰	۱,۴۷	۱,۳	آموزشی
۰,۸۰	۰,۱۳	۱,۰۳	۰,۱۴	۱,۱۷	۰,۲۴	۰,۴۴	۰,۱۲	۰,۱۴	۰,۱۶	۱,۳	بهداشتی- درمانی
۰,۲۰	۲,۹۴	۱,۸۰	۴,۶۱	۰,۳۶	۱,۷۲	۴,۹۹	۱,۷۸	۰,۱۲	۲,۲۶	-	پارک و فضای سبز عمومی
۰,۳۱	۰,۴۳	۰,۲۹	۰,۵۷	۰,۵۴	۰,۲۹	۰,۶۶	۰,۵۵	۰,۶۷	۰,۰۸	۲,۴	ورزشی
۰,۹۸	۰,۴۱	۲,۳۳	۰,۲۵	۰,۴۱	۰,۶۷	۰,۳۴	۰,۳۳	۰,۴۱	۰,۵۵	۰,۸	مذهبی- فرهنگی
۰,۶۶	۰,۲۰	۱,۸۲	۱,۴۲	۲,۷۸	۱,۳۳	۰,۶۳	۰,۳۱	۰,۰۸	۰,۲۱	۰,۴	اداری و انتظامی
۰,۶۱	۱,۰۰	۱,۱۸	۰,۲۴	۰,۵۲	۰,۹۸	۰,۷۸	۱,۶۴	۰,۵۳	۰,۳۵	-	حمل و نقل و انبارداری
۱,۲۱	۰,۲۳	۰,۵۷	۰,۱۷	۱,۴۳	۴,۵۹	۰,۱۳	۰,۰۹	۰,۱۱	۰,۱۲	۰,۴	تاسیسات و تجهیزات شهری
۰,۶۸	۲۶,۷۰	۰,۱۰	۴,۰۲	۱,۰۹	۰,۰۵	۱۳,۱۵	۱۰,۵۲	۱۱,۹۳	۶,۱۰	-	اراضی کشاورزی
۲۲,۴۴	۲۰,۶۴	۲۱,۷۹	۲۵,۴۷	۱۶,۳۹	۱۴,۴۹	۲۲,۰۲	۲۱,۲۰	۲۰,۵۹	۲۴,۳۶	۳۴,۸	شبکه های ارتباطی و فضای باز
۱,۷۰	۳,۸۰	۱,۱۲	۱,۴۶	۵,۳۳	۱,۱۱	۱,۷۰	۹,۳۷	۱,۲۳	۹,۶۹	-	صنعتی و کارگاهی
۱۶,۳۵	۱۴	۱۵,۷۷	۲۰,۹۲	۴۱,۷۸	۴۶,۶۱	۲۴,۶۲	۲۴,۴۶	۱۶,۶۳	۱۴,۳۶	۱۳,۴	سایر

جدول ۴-۴ سطح نواحی اصفهان و کاربری اراضی بر اساس هکتار

ردیف	کاربری	منطقه ۱		منطقه ۳		حوزه مرکز	
		سطح (هکتار)	نسبم (درصد)	سطح (هکتار)	نسبم (درصد)	سطح (هکتار)	نسبم (درصد)
۱	مسکونی	۵۰۹/۱۸	۴۸/۸۴	۵۳۱/۸۵	۴۴/۹۸	۴۷/۵۳	۴۶/۷۹
۲	مختلط تجاری	۳۹/۹۶	۳/۸۳	۶۲/۴۸	۵/۲۸	۱۰۲/۴۴	۴/۶۰
۳	آموزشی	۱۴/۵۲	۱/۳۹	۲۲/۹۰	۱/۹۴	۳۷/۴۲	۱/۶۸
۴	بهداشتی و درمانی	۸/۳۲	۰/۸۰	۱۲/۳۵	۱/۰۴	۲۰/۶۷	۰/۹۳
۵	پارک و فضای سبز عمومی	۲/۰۹	۰/۲۰	۲۱/۲۳	۱/۸۰	۲۳/۳۲	۱/۰۵
۶	پارک طبیعی حاشیه زاینده رود	۱۵/۸۷	۱/۵۲	۱۱/۷۳	۰/۹۹	۲۷/۶۰	۱/۲۴
۷	ورزشی	۳/۲۸	۰/۳۱	۳/۴۰	۰/۲۹	۶/۶۸	۰/۳۰
۸	فرهنگی	۲/۵۱	۰/۲۴	۸/۱۰	۰/۶۸	۱۰/۶۱	۰/۴۸
۹	مذهبی	۷/۶۸	۰/۷۴	۱۹/۵۶	۱/۶۵	۲۷/۲۳	۱/۲۲
۱۰	گردشگری و پذیرایی	۱/۳۱	۰/۱۳	۲/۴۸	۰/۲۱	۳/۸۰	۰/۱۷
۱۱	مجموعه فرهنگی، مذهبی، تفریح فولاد	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
۱۲	اداری و انتظامی	۶/۸۷	۰/۶۶	۲۱/۴۹	۱/۸۲	۲۸/۳۵	۱/۲۷
۱۳	خدمات اجتماعی	۱/۲۸	۰/۱۲	۱/۰۷	۰/۰۹	۲/۳۴	۰/۱۱
۱۴	آموزش عالی و فنی و حرفه ای	۵/۸۶	۰/۵۶	۳/۷۳	۰/۳۲	۹/۵۹	۰/۴۳
۱۵	حمل و نقل و بار داری	۶/۳۲	۰/۶۱	۱۳/۹۶	۱/۱۸	۲۰/۲۸	۰/۹۱
۱۶	تاسیسات و تجهیزات شهری	۱/۲۱	۰/۱۲	۶/۷۸	۰/۵۷	۷/۹۹	۰/۳۶
۱۷	صنعتی و کارگاهی	۱۷/۷۶	۱/۷۰	۱۳/۲۲	۱/۱۲	۳۰/۹۷	۱/۳۹
۱۸	نظامی	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
۱۹	اراضی کشاورزی	۷/۰۸	۰/۶۸	۱/۲۲	۰/۰۹	۸/۳۰	۰/۳۷
۲۰	باغ و پرورش گل گیاه	۱/۳۲	۰/۱۳	۰/۳۷	۰/۰۳	۱/۶۹	۰/۰۸
۲۱	پارک گوهستان	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
۲۲	درخت کار و جنگل کاری	۰/۰۸	۰/۰۱	۰/۰۱	۰/۰۰	۰/۰۸	۰/۰۰
۲۳	حرایم و میانه های سبز معابر	۱۶/۳۶	۱/۵۶	۲۴/۴۳	۲/۰۷	۴۰/۶۹	۱/۸۳
۲۴	ارتفاعات	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
۲۵	مجازی آبی	۱۷/۹۹	۱/۷۳	۱۸/۴۴	۱/۵۶	۳۶/۴۳	۱/۶۴
۲۶	شبکه ارتباطی و فضای باز	۲۳۳/۹۹	۲۲/۴۴	۲۵۷/۶۳	۲۱/۷۹	۴۹۱/۶۲	۲۲/۱۰
۲۷	بایر و بدین استفاده	۱۲۱/۸۵	۱۱/۶۹	۱۲۴/۰۲	۱۰/۴۹	۲۴۵/۸۶	۱۱/۰۸
۲۸	سایر استفاده ها	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
۲۹	گوی شهید کشوری	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰	۰/۰۰
جمع		۱۰۴۲/۵۷	۱۰۰/۰۰	۱۴۰/۳۹	۱۱۸۲/۴۳	۱۰۵/۶۷	۲۲۲۵/۰۰
جمعیت		۷۴۲۶۳		۱۱۱۹۰۰		۱۸۶۱۶۳	
تراکم تکالض		۱۴۵/۸۵		۲۱۰/۴۰		۱۷۸/۸۳	
تراکم ناخالص		۷۱/۲۳		۹۴/۶۴		۸۳/۶۷	

۳-۴ آماده سازی داده ها

در این بخش به بررسی نحوه پالایش و آماده سازی داده ها جهت استفاده در شبکه عصبی

می پردازیم.

۴-۳-۱ طبقه‌بندی کاربری اراضی

فایل های مرتبط با کاربری اراضی تهیه شده از سازمان فاوا شهرداری مشهد و شهرداری اصفهان، شامل کاربری های متنوعی بودند. از آنجا که تولید و جذب سفر به هدف سفرها مرتبط اند، کاربری ها را بر اساس اهداف سفر آنها در دسته های اداری، آموزشی، تفریحی، درمانی، حمل و نقل، تجاری، صنعتی، مسکونی و خطا طبقه بندی شدند.

هر طبقه شامل دسته های زیر بود:

اداری:

اداری - انتظامی، اداری - دولتی، تاسیسات و تجهیزات شهری، تاسیسات-نظامی-انتظامی، تجهیزات شهری، ذخیره خدمات، تجهیزات شهری (پمپ بنزین یا آتش نشانی)، خدمات (برای انبارها و کارگاهها)، مختلط - اداری و آموزشی، کاربری ویژه (اداری انتظامی مسکن استجار و مسکون)، تجاری ۱٪۲۰۰، تجاری تراکم ۱٪۲۵۰، تجاری محله، تاسیسات و تجهیزات شهری - ایستگاه تقلیل فشار، خدمات اجتماعی.

آموزشی:

آموزشی - دبستان، آموزشی - متوسطه اول، ، آموزشی - متوسطه دوم، آموزشی - مهد کودک، آموزشی - هنرستان، آموزشی، آموزش - آموزش عالی، دبستان، آموزشی - متوسطه اول، آموزشی - متوسطه دوم، آموزشی - مدرسه علمیه، آموزشی - مهد کودک، آموزشی - ویژه، آموزشی - هنرستان، کشاورزی - آموزشی، آموزش - هنرستان کشاورزی.

تفریحی:

خدمات رفاهی- تربیتی، فضای سبز، فرهنگی - تاریخی، فرهنگی - مذهبی، مختلط - بازی و فضای تفریحی و فضای سبز، ورزشی، ورزشی - زمین ورزشی و بازی، اقامتی ویژه، باغ، توریستی - تفریحی، توریستی - مسافرخانه، توریستی - ویژه، توریستی - هتل، فضای سبز - حفاظتی، مختلط

- بازی و فضای تفریحی و فضای سبز، مختلط - بازی و فضای تفریحی و مسیر پیاده، مجموعه تفریحی - تاریخی - فرهنگی باغ خونی، مختلط - اقامتی با خدمات گردشگری، فرهنگی - مجموعه فرهنگی یادمانی، ورزشی، ورزشی - مجتمع ورزشی، ورزشی - مقیاس ناحیه ای، فضای سبز - تجهیز شده.

درمانی:

بهداشتی - درمانی، درمانی موجود، درمانی ناحیه ای، سرویسهای بهداشتی.

حمل و نقل:

پارکینگ، حمل و نقل - پایانه باربری، مختلط - پایانه مسافربری و انبارداری و پارکینگ، انبارداری - پارکینگ و حمل و نقل، خدمات شهری - پمپ بنزین، حمل و نقل - پایانه بار کامیون داران، مختلط - حمل و نقل و انبارداری.

تجاری:

بازار روز، تجاری، تجاری محله، تجاری ناحیه، تراکم مختلط تجاری، اداری، مسکونی، تجاری - مسکونی ۵۰/۵۰، تجاری ۵، تجاری اقامتی ۲۰-۸۰، تجاری مسکونی ۱۰-۹۰، تجاری مسکونی ۲۰-۸۰، تجاری مسکونی ۳۰-۷۰، راسته بازارچه باغ خونی، اداری - تجاری، تجاری محله، تجاری مختلط، تجاری مختلط (مرتبط با بخش صنعتی سبک)، تجاری مقیاس ناحیه، تجاری، مسکونی ناحیه ای، تجاریهای میوه و تره بار، مختلط - تجاری مسکونی، مختلط - تجاری و کارگاهی و تعمیرگاهی، نمایشگاهی.

صنعتی:

صنعتی، صنعتی - صنایع سبک، صنعتی - صنایع کشاورزی، صنعتی - صنایع سبک و کارگاهی، انبار داری و کارگاههای غیر مزاحم، صنعتی - صنایع سنگین، صنعتی - کارگاهی، صنعتی - مجتمع های تعمیرگاهی.

مسکونی:

مسکونی موجود، مسکونی با تراکم بسیار زیاد، مسکونی با تراکم زیاد، مسکونی با تراکم کم، مسکونی موجود، مسکونی، مسکونی تراکم متوسط، مسکونی تراکم متوسط (آپارتمانی)، مجتمع ۱۲۰٪، مجتمع آپارتمانی، ساختمانهای مسکونی مرتفع ۱۸ تا ۲۲ طبقه، ساختمانهای مسکونی ویژه، شهرک های موجود و مصوب.

خطا:

فیلد بدون نام، کشاورزی - اراضی کشاورزی، مسکونی پیشنهادی، فضای سبز - حریم، کال چهل بازه، کال، کشاورزی - باغهای زمستانی، مختلط، ذخیره، باغ شخصی، باغ شخصی و گاوداری، ذخیره، راه سواره، زمین ذخیره، واگذاری مسکونی، مجاری آبی، شبکه ارتباطی و فضای باز، بایر و بدون استفاده، سایر استفاده ها، کوی شهید کشوری.

۴-۳-۲ طبقه بندی سرشماری جمعیتی

داده های مربوط به سرشماری در گروه های

جمعیت کل، جمعیت مرد، جمعیت زن، تعداد کل خانوار، جمعیت ۶ ساله و بیشتر، جمعیت ۶ ساله و بیشتر با سواد، جمعیت ۶ ساله و بیشتری سواد، مرد ۶ ساله و بیشتر، جمعیت ۱۰ ساله و بیشتر، جمعیت ۱۰ ساله و بیشتر مرد، جمعیت ۱۰ ساله و بیشتر زن، جمعیت شاغل مرد، جمعیت شاغل زن، جمعیت بیکار مرد، جمعیت بیکار زن، وضع فعالیت کد ۵ (محصل)، وضع فعالیت کد ۶ (خانه دار)، وضع فعالیت کد ۷ (دارای درآمد بدون کار)، وضع فعالیت کد ۸ (سایر)

طبقه بندی شدند.

۳-۳-۴ آماده سازی داده های شهر مشهد

• ماتریس سفر

ماتریس سفر شهر مشهد مربوط به داده‌های سال ۱۳۹۵ در ابتدا بر اساس نواحی ترافیکی بود.

لذا با توجه به

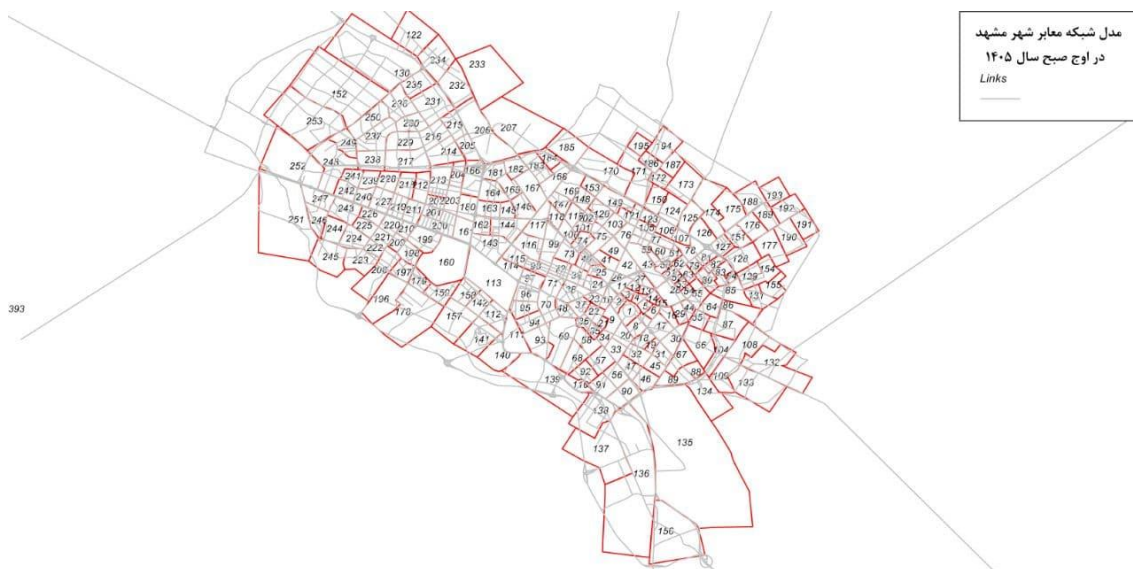
۱. نقشه ناحیه بندی نواحی ترافیکی مشهد (شکل ۴-۱) که شامل ۲۵۳ ناحیه می‌باشد.

۲. نقشه ناحیه بندی شهرداری مشهد (شکل ۴-۲) که شامل ۱۳ ناحیه می‌باشد.

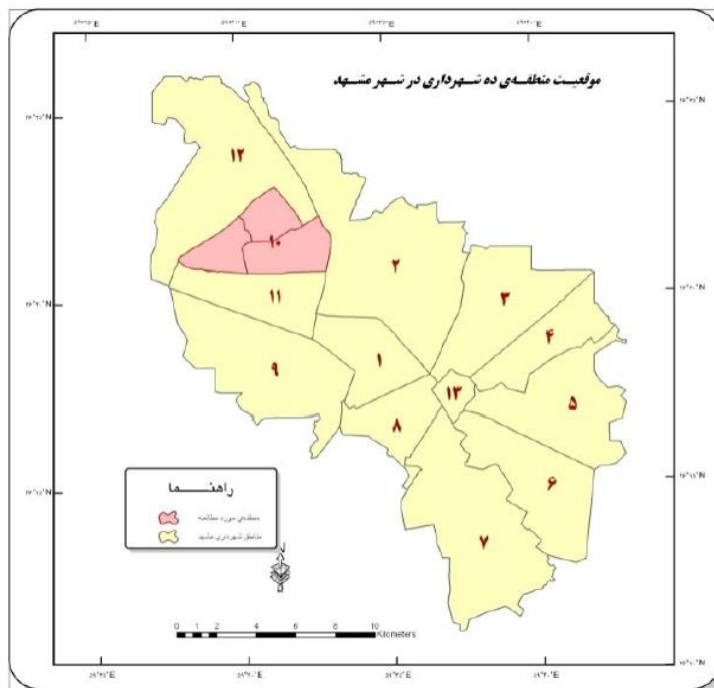
۳. و با استفاده از نرم افزار Photoshop

ماتریس سفر بر اساس نواحی شهرداری محاسبه شد. بدین صورت که ابتدا تصاویر ناحیه‌بندی ترافیکی و شهرداری بوسیله نرم افزار فتوشاپ بر یکدیگر منطبق شدند و ملاحظه شد که هر کدام از نواحی ترافیکی در کدام ناحیه شهرداری واقع شده است. سپس ماتریس سفر اولیه که بر اساس نواحی ترافیکی بود، بر اساس نواحی شهرداری تهیه

شد (جدول ۴-۸).



شکل ۴-۱ ناحیه بندی شهر مشهد بر اساس نواحی ۲۵۳ گانه ترافیکی



شکل ۴-۲ ناحیه بندی شهر مشهد بر اساس نواحی سیزده گانه ی شهرداری

جدول ۴-۵ ماتریس نهایی سفر مشهد بر اساس نواحی شهرداری

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	111337.9	58343.69	18357.45	3769.531	2457.581	4593.683	14702.05	29058.39	27857.52	13761.38	19834.57	4745.723	9950.148
2	55445.06	149592.8	48432.51	9065.397	3837.867	4695.815	8867.236	18692.46	21630.47	33106.38	26661.3	3186.527	12158.77
3	20316.09	49557.6	139789.3	44341.41	10004.78	11936.43	12360.8	14950.45	9188.542	11825.7	9115.957	1338.852	19930.03
4	4367.845	10394.28	44688.39	92851.65	14909.18	14376.82	9799.127	5883.19	2189.383	2524.265	1852.237	274.4003	8426.789
5	2532.553	4347.432	10443.24	15004.91	21256.1	12248.31	7600.979	3563.435	1406.24	1056.653	914.3023	118.8098	5456.175
6	5729.416	5973.76	13152.72	14570.67	12653.49	65638.62	24634.55	11353.26	3977.476	1785.101	197666	205.0642	12403.45
7	15882	10514.2	12838.29	10118.86	8007.324	22488.45	106923.4	33297.89	10603.9	4439.288	6878.525	615.5871	14949.21
8	31516.18	19942.08	18488.81	7228.695	4041.517	9410.718	33052.56	55222.41	16026.2	8573.058	7395.27	1099.813	13746.74
9	38043.16	22767.87	10241.77	2553.403	2320.248	4445.271	15401.5	23651.7	154096.7	20226.33	49086.52	6405.873	4804.742
10	17407.57	34695.78	13981.54	2855.835	1081.416	1618.986	3945.32	8798.3	19817.02	156721.9	35327.91	12868.93	3777.949
11	23428.75	24283.83	8369.818	1761.822	754.039	1520.389	5172.285	10203.26	40936.62	27672.26	88861.41	4237.591	3328.186
12	1994.888	2865.429	1429.36	269.4181	115.5312	182.741	572.1462	1058.389	4239.863	9531.414	4221.25	6637.372	418.1154
13	10362.66	13233.96	22341.23	9305.875	5358.833	10648.98	16193.18	15707.78	4065.898	3990.787	3439.689	400.5642	19860.79

• سرشماری جمعیت

ابتدا پارامترهای نامرتب مانند تعداد منازل استیجاری حذف شدند و سپس داده‌ها جمع بندی

و آماده استفاده شدند. داده‌های نهایی استفاده شده در این تحقیق در جدول ۴-۹ درج گردید.

جدول ۴-۶ داده‌های آماری سرشماری شهر مشهد سال ۱۳۹۵

تامن	12	11	10	9	8	7	6	5	4	3	2	1	مناطق شهرداری
13211	83159	199969	295674	325986	92071	252633	231908	175084	261846	381765	511592	166450	جمعیت کل
6717	42091	102335	148292	161239	45674	127754	116351	86876	131298	192536	256192	81320	جمعیت مرد
6494	41068	97634	147382	164747	46397	124879	115557	88208	130548	189229	255400	85130	جمعیت زن
4357	26383	60550	91690	102860	29863	76316	66845	48445	76599	115481	157161	55053	تعداد کل خانوار
12060	72784	185976	264446	298725	84216	220406	199809	152033	227947	331376	450692	155420	جمعیت ۶ ساله و بیشتر
10697	70922	181066	254711	290226	79350	204747	171397	131546	203312	301328	421528	150276	جمعیت ۶ ساله و بیشتر باسواد
1359	1855	4754	9703	8253	4860	15651	28398	20478	24608	30023	26769	5000	جمعیت ۶ ساله و بیشتری سواد
6139	36725	95276	132325	147283	41638	111273	99810	75094	113916	166726	224939	75739	مرد ۶ ساله و بیشتر
11351	67268	177259	247020	282285	79196	202206	181697	139167	208717	303413	415938	148040	جمعیت ۱۰ ساله و بیشتر
5771	33843	90839	123443	138829	39076	101996	90448	68619	103980	152292	207241	72023	جمعیت ۱۰ ساله و بیشتر مرد
5580	33425	86420	123577	143456	40120	100210	91249	70548	104737	151121	208697	76017	جمعیت ۱۰ ساله و بیشتر زن
3143	21035	40763	67073	71512	21016	64466	56912	44930	67290	96678	125602	36678	جمعیت شاغل مرد
644	4099	13691	14440	22770	4569	8018	9016	6831	7427	11389	21365	10844	جمعیت شاغل زن
319	2284	5623	9549	9285	2128	5616	7100	4649	6201	9873	13638	3971	جمعیت بیکار مرد
85	1205	3328	4399	6090	904	1234	943	1172	1317	1817	5099	2264	جمعیت بیکار زن
2206	13152	37294	49413	66300	16089	35899	28852	24494	36507	51891	78532	31989	وضع فعالیت در ۵ (محصّل)
3550	20166	42422	74211	70847	23891	69154	63390	47430	73811	106522	133051	39271	وضع فعالیت در ۶ (خانه دار)
733	3604	19435	21132	26659	7487	11437	8455	3674	7008	13214	23320	16703	وضع فعالیت در ۷ (دارای درآمد بدون کار)
660	1709	5601	6748	8699	3100	6360	7004	5959	9112	11992	12974	6141	وضع فعالیت در ۸ (سایر)

• کاربری اراضی

ابتدا داده ها باتوجه به دسته بندی‌های تعیین شده در بخش ۴-۳-۱ طبقه بندی شدند و

سپس باتوجه به شماره نواحی تفکیک شده و در نهایت جمع بندی و آماده استفاده شدند(جدول ۴-۱۰).

جدول ۴-۷ سطح زیربنای کاربری اراضی داخل هر ناحیه شهرداری شهر مشهد در سال ۱۳۹۵

zone	1	2	3	4	5	6	7	8	9	10	11	12
اداری	354990.9	504386.3	337824.8	69632.66	38154.49	310853.8	2373681	2410259	2209654	171938.7	252665.7	102623.4
آموزشی	768558.5	1003207	792865.9	480738.3	355560.7	591072.4	6231470	1010783	5965836	864775.1	499891.4	1103180
تفریحی	868674	4014618	3546812	1163585	1844303	2340361	10125338	4254118	13575187	1938401	2863577	10080650
درمائی	187923.8	322371.9	232529.3	222310.5	126813.4	254140.6	222357	549374.7	258296.6	167964.5	180814.4	256270.3
حمل و نقل	3793.175	432355.7	1268422	1075.749	458808.4	163921	6652841	491201.1	160019	655019.1	0	471450.3
تجاری	691581	1392679	655743.6	59199.48	365263.8	365648.2	1014733	1029596	651486.7	323737.6	383718.3	2048058
صنعتی	5177.982	675229.8	169188.5	20723.61	1945617	979861.7	5519590	7789.044	127663.5	16996.83	0	582022
مسکونی	8473277	12451435	9764308	5624028	4082675	6634765	13036274	6964089	17819824	9412530	6773594	5864631

۴-۳-۴ آماده سازی داده های شهر اصفهان

• ماتریس سفر

ناحیه بندی شهر اصفهان بگونه ای است که تعدادی از نواحی با یکدیگر تلاقی دارند(شکل ۴-۳).



شکل ۴-۳ ناحیه‌بندی شهرداری اصفهان

برای مثال به سمت غرب اصفهان می‌توان اشاره کرد که ناحیه ۱۱ شهرداری از وسط نواحیه ۲ عبور کرده است. در داده‌های آماری مربوط به کاربری اراضی، مجموع سطح زیربنایی موجود در منطقه ۱۱ و ۲ در مجموع بیان شده بود و تفکیک شده نبود (جدول ۴-۱۱). لذا برای انجام محاسبات باید ماتریس سفر را نیز بر همین اساس اصلاح کرد. در این تحقیق نواحی‌ای که با یکدیگر اشتراک داشتند، بعنوان یک ناحیه در نظر گرفته شدند و بدین صورت اصفهان به ۱۰

ناحیه شهری تقسیم بندی شد (جدول ۴-۱۲)

جدول ۴-۸ ماتریس سفر اصلی شهر اصفهان مربوط به سال ۱۳۹۵

zone	1	2 & 11	3	4	5 & 13	6	7 & 14	8 & 12	9	10
1	80043	26457	34120	9205	31884	15579	24955	60083	27972	12286
2 & 11	27649	108740	10298	2526	10058	5992	7316	54154	13972	2745
3	33309	9994	128391	57810	32509	36283	66456	46817	7715	65048
4	9412	2570	59776	106888	18269	19752	11458	10189	1504	22946
5 & 13	33575	9534	35284	18490	447079	99441	21632	33799	9432	13521
6	16211	6071	37019	18917	99471	140845	18475	20739	4351	14511
7 & 14	26024	7390	69969	11695	23401	18468	237954	66206	2312	58204
8 & 12	62734	54545	50440	10889	35321	21479	65778	409513	8519	19368
9	28553	14140	7569	1564	9150	4099	2122	8330	46746	1694
10	12405	2659	68755	22983	13118	15124	55824	18912	1641	151574

جدول ۴-۹ ماتریس سفر استفاده شده در این تحقیق، شهر اصفهان مربوط به سال ۱۳۹۵

new zone		1	2	3	4	5	6	7	8	9	10
	zone	1	2 & 11	3	4	5 & 13	6	7 & 14	8 & 12	9	10
1	1	80043	26457	34120	9205	31884	15579	24955	60083	27972	12286
2	2 & 11	27649	108740	10298	2526	10058	5992	7316	54154	13972	2745
3	3	33309	9994	128391	57810	32509	36283	66456	46817	7715	65048
4	4	9412	2570	59776	106888	18269	19752	11458	10189	1504	22946
5	5 & 13	33575	9534	35284	18490	447079	99441	21632	33799	9432	13521
6	6	16211	6071	37019	18917	99471	140845	18475	20739	4351	14511
7	7 & 14	26024	7390	69969	11695	23401	18468	237954	66206	2312	58204
8	8 & 12	62734	54545	50440	10889	35321	21479	65778	409513	8519	19368
9	9	28553	14140	7569	1564	9150	4099	2122	8330	46746	1694
10	10	12405	2659	68755	22983	13118	15124	55824	18912	1641	151574

• سرشماری جمعیتی

داده های مربوط به سرشماری جمعیتی شهر اصفهان مربوط به ۱۵ ناحیه به تفکیک نواحی بودند. لذا باتوجه به ناحیه بندی جدید مورد استفاده در این تحقیق (۱۰ ناحیه) داده های جمعیتی دسته بندی و آماده تحلیل شدند (جدول ۴-۱۳).

جدول ۴-۱۰ داده‌های سرشماری جمعیتی سال ۱۳۹۵ شهر اصفهان

مناطق شهرداری	1	2	3	4	5	6	7	8	9	10
جمعیت کل	79091	127925	110368	133076.5	278658.5	112042.8	329053.7	374144.1	72252.57	207688
جمعیت مرد	38589	64618	54008	65917	142181	54776	166521	187361	36186	104743
جمعیت زن	40214	62283	55609	66041	135891	57004	161683	185346	35362	102890
تعداد کل خانوار	27187	39436	36791	43277	85217	37509	99414	117055	23217	65571
جمعیت ۶ ساله و بیشتر	73891	114397	102271	122991	261610	105400	295505	340775	65443	188748
جمعیت ۶ ساله و بیشتر باسواد	70023	101158	95587	115826	251466	100337	266602	318022	59244	171876
جمعیت ۶ ساله و بیشتری سواد	3860	13228	6623	7155	9777	5052	28895	22724	6197	16862
مرد ۶ ساله و بیشتر	36076	58165	50293	61210	133847	51529	149747	170897	33109	95090
جمعیت ۱۰ ساله و بیشتر	70761	106770	97469	117060	250333	100814	274958	320292	61815	177091
جمعیت ۱۰ ساله و بیشتر مرد	34420	54299	47839	58198	127951	49200	139224	160350	31220	89107
جمعیت ۱۰ ساله و بیشتر زن	36341	52471	49630	58862	122382	51614	135734	159942	30595	87984
جمعیت شاغل مرد	18179	33542	24969	31056	57107	24977	83735	91848	18711	51974
جمعیت شاغل زن	4591	3493	5808	7485	17139	7770	11559	17441	2563	8046
جمعیت بیکار مرد	2301	3962	3089	4107	9225	3398	10020	12357	1980	5930
جمعیت بیکار زن	1359	962	1557	1998	6417	2416	3199	5633	764	2134
وضع فعالیت کد ۵ (محصل)	13824	18738	19587	22735	53632	19733	52035	60529	10747	32879
وضع فعالیت کد ۶ (خانه دار)	20790	36146	28743	34046	61184	26222	89288	97478	20367	57238
وضع فعالیت کد ۷ (دارای درآمد بدون کار)	7537	7269	10752	12419	28321	13408	18583	27122	4858	14458
وضع فعالیت کد ۸ (سایر)	2170	2635	2918	3189	5830	2872	6530	7816	1821	4401

• کاربری اراضی

ابتدا با توجه به جداول ۴-۶ و ۴-۷ مساحت هر منطقه ۱۵ گانه شهرداری تعیین شد. سپس

مساحت بدست آمده برای هر منطقه شهرداری ضربدر سهم هریک از کاربری‌های موجود در آن

منطقه شد تا زیربنای هر کاربری در هر کدام از مناطق بدست آید (جدول ۴-۱۴).

جدول ۴-۱۱ سطح زیربنای کاربری اراضی مختلف (بر اساس هکتار) در شهر اصفهان در سال ۱۳۹۵

zone	1	2 & 11	3	4	5 & 13	6	7 & 14	8 & 12	9	10
اداری	9.36	7.47	29.34	16.95	607.57	465.5	19.21	20.65	2.11	5.3
آموزشی	20.38	13.76	26.63	14.4	277.54	35.57	28.15	28.94	8.67	23.54
تفریحی	34.14	68.41	66.87	146.42	355.24	148.27	217.8	125.65	182.41	49.64
درمانی	8.32	2.73	12.35	1.46	35.9	4.08	10.57	4.12	1.51	2.55
حمل و نقل	6.32	17	13.96	2.54	15.86	16.45	18.73	54.16	5.69	5.65
تجاری	39.96	28.49	62.48	10.66	26.71	19.94	25.46	32.36	17.91	20.95
صنعتی	17.76	64.42	13.22	15.55	164.05	18.72	40.75	310.13	13.28	155.47
مسکونی	509.18	459.36	531.85	409.51	815.62	431.14	685.29	921.92	493.01	625.17

۴-۳-۵ ضریب همبستگی میان داده ها

• تعریف

ضریب همبستگی ابزاری آماری برای تعیین نوع و درجه‌ی رابطه‌ی یک متغیر کمی با متغیر

کمی دیگر است. ضریب همبستگی، یکی از معیارهای مورد استفاده در تعیین همبستگی دو متغیر

است. ضریب همبستگی شدت رابطه و همچنین نوع رابطه (مستقیم یا معکوس) را نشان می‌دهد.

این ضریب بین ۱ تا ۱- است و در صورت عدم وجود رابطه بین دو متغیر، برابر صفر است.

همبستگی بین دو متغیر تصادفی X و Y به صورت زیر تعریف می‌شود:

$$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \quad \text{معادله ۴-۱}$$

که در آن E عملگر امید ریاضی، cov به معنای کوواریانس، corr نماد معمول برای همبستگی پیرسون، و سیگما نماد انحراف معیار است.

احتمالا، گسترده ترین کاربرد شاخص آماری همبستگی دو متغیری، ضریب همبستگی گشتاوری پیرسون است که به طور معمول همبستگی پیرسون نامیده میشود. ضریب پیرسون نشان می دهد که تا چه اندازه بین متغیرهای کمی رابطه خطی وجود دارد.

کاربرد اصلی ضریب پیرسون زمانی است که متغیرها از نوع پارامتری باشند، بدین معنا که توزیع نرمال داشته باشند و در سطح فاصله ای/نسبی باشند. البته زمانی که متغیرها از نوع شبه فاصله ای باشند (یعنی هر متغیر ترکیبی از چند متغیر ترتیبی باشد که اصطلاحا به آن مقیاس های تراکمی میگویند)، برخی از پژوهشگران از ضریب پیرسون استفاده میکنند. برخی از نویسندگان استفاده از ضریب پیرسون برای یک متغیر دو ارزشی و یک متغیر فاصله ای/نسبی را هم مجاز شمرده اند. تفسیر همبستگی پیرسون زمانی که یکی از متغیرها دوارزشی (فقط شامل دو سطح) اما متغیر دیگر کمی است نیز میتواند منطقی باشد.

• تفسیر شدت رابطه در همبستگی پیرسون

بعد از تعیین معنی داری و جهت رابطه، باید شدت رابطه ارزیابی شود. برای تفسیر شدت رابطه دومتغیر، تقسیم بندی های گوناگونی ارائه شده است. تقسیم بندی ارائه شده در جدول ۴-۱۵ یکی از تقسیم بندی های متداول است.

جدول ۴-۱۲ شیوه تفسیر شدت رابطه در همبستگی پیرسون

شدت رابطه	تفسیر
۰/۸ تا ۱	رابطه بسیار قوی
۰/۶ تا ۰/۸	رابطه قوی
۰/۴ تا ۰/۶	رابطه متوسط
۰/۲ تا ۰/۴	رابطه کم (یا ضعیف)
صفر تا ۰/۲	فقدان رابطه یا رابطه ناچیز

۴-۳-۶ بررسی ضریب همبستگی پارامترها در نرم افزار SPSS

داده ها بوسیله نرم افزار spss version21 64bit جهت انجام محاسبات مربوط به ضریب

همبستگی پیرسون به کار گرفته شدند.

برای تعیین همبستگی در نرم افزار spss مسیر زیر دنبال می شود:

Analyze--->Correlate--->Bivariate

فایل خروجی بصورت ۲ فایل excel که هر کدام شامل ۵۶ ستون و ۲۰۲ ردیف بودند و با نام((همبستگی spss مشهد)) و ((همبستگی spss اصفهان)) در CD ضمیمه تحقیق آورده شده اند.

برای مثال بخشی از جدول مربوط به داده های شهر مشهد را بررسی می کنیم. در این بخش(جدول ۴-۱۶) از جدول بدست آمده(جدول ۴-۱۷) ابتدا باید به سطر دوم دقت نمود که در جدول با Sig. (2-tailed) نشان داده شده است و نمایانگر p-value یا همان وجود یا عدم وجود رابطه معنادار میان دو پارامتر است. اگر این مقدار کمتر از ۰,۰۵ باشد، رابطه معناداری میان دو پارامتر وجود دارد و سپس باتوجه به ضریب پیرسون که در جدول با Pearson Correlation نشان

داده شده است، مقدار همبستگی آن دو پارامتر را میتوان تعیین کرد. اگر مقداری منفی باشد یعنی با زیاد شدن یکی، دیگری کاهش خواهد یافت (رابطه معکوس) و اگر مثبت باشد، یعنی با افزایش یکی، دیگری نیز افزایش خواهد یافت (رابطه مستقیم) و اگر صفر باشد، یعنی آن دو پارامتر از یکدیگر مستقل هستند و با افزایش یا کاهش یکی از پارامترها، پارامتر مقابل، تغییری نخواهد کرد. با توجه به جدول ۴-۱۶ میتوان ابتدا باتوجه به مقادیر سطر دوم بیان کرد که میان داده‌های کاربری اداری ناحیه مبداء و داده‌های کاربری اداری مقصد از آنجا که مقدار $p\text{-value}$ برابر ۱ شده است و از ۰,۰۵ بیشتر است میان این دو پارامتر رابطه معنا داری وجود ندارد. همچنین مشاهده می‌شود که مقدار ضریب پواسون نیز در این حالت برابر ۰ است که نشان‌دهنده رابطه مستقل این دو پارامتر از یکدیگر است. اما برای پارامترهای کاربری اداری مبداء و کاربری آموزشی مبداء مقدار $p\text{-value}$ برابر ۰ شده است که نشانگر آن است که میان داده‌های این دو پارامتر رابطه معناداری وجود دارد و از آنجا که ضریب همبستگی پیرسون میان این دو برابر با ۰۰۶/۸ است مشخص می‌شود که با یکدیگر رابطه مستقیم و نسبتاً قوی ای دارند. در تحقیق هنگامی که پارامترهای با ضریب همبستگی بالای ۰/۸ بررسی شوند، باید ملاحظه کرد میان این دو پارامتر کدام یک واریانس داده‌های بالاتری دارد و از آن پارامتر در محاسبات استفاده نمود. اما از آنجا که بررسی این مقادیر بسیار زمانبر است، در این تحقیق بوسیله کد برنامه نویسی شده نرم افزار متلب که قادر است ضریب همبستگی مورد نظر را میان داده‌ها بررسی کند و در میان پارامترهای همبسته، پارامتر دارای واریانس بیشتر را معرفی کند استفاده شد. توضیحات این کد در بخش ۴-۳-۸ آورده شده است و نتایج بررسی ضرایب همبستگی در بخش ۵-۲-۱ و ۵-۲-۲ آورده شدند.

جدول ۴-۱۳ دو سطر و پنج ستون از خروجی نرم افزار SPSS پارامترهای شهر اصفهان

		اداری_مبداء	اداری_مقصد	آموزشی_مبداء	آموزشی_مقصد
اداری_مبداء	Pearson Correlation	1	.000	.806**	.000
	Sig. (2-tailed)		1.000	.000	1.000
	N	100	100	100	100

جدول ۴-۱۴ بخشی از جدول همبستگی های خروجی نرم افزار SPSS داده های شهر اصفهان

		اداری_مبداء	اداری_مقصد	آموزشی_مبداء	آموزشی_مقصد	تفریحی_مبداء	تفریحی_مقصد	ترجمتی_مبداء	ترجمتی_مقصد	حملاتقل_مبداء	حملاتقل_مقصد	تجاری_مبداء	تجاری_مقصد	منتقلی_مبداء
اداری_مبداء	Pearson Correlation	1	.000	.806	.000	.682	.000	.699	.000	.032	0.000	-.148	.000	.108
	Sig. (2-tailed)		1.000	.000	1.000	.000	1.000	.000	1.000	.752	1.000	.140	1.000	.285
	N	100	100	100	100	100	100	100	100	100	100	100	100	100
اداری_مقصد	Pearson Correlation	.000	1	.000	.806	.000	.682	.000	.699	.000	.032	.000	-.148	.000
	Sig. (2-tailed)	1.000		1.000	.000	1.000	.000	1.000	.000	1.000	.752	1.000	.140	1.000
	N	100	100	100	100	100	100	100	100	100	100	100	100	100
آموزشی_مبداء	Pearson Correlation	.806	.000	1	.000	.782	.000	.942	.000	.053	.000	-.015	0.000	.319
	Sig. (2-tailed)	.000	1.000		1.000	.000	1.000	.000	1.000	.602	1.000	.881	1.000	.001
	N	100	100	100	100	100	100	100	100	100	100	100	100	100
آموزشی_مقصد	Pearson Correlation	.000	.806	.000	1	.000	.782	.000	.942	.000	.053	.000	-.015	.000
	Sig. (2-tailed)	1.000	.000	1.000		1.000	.000	1.000	.000	1.000	.602	1.000	.881	1.000
	N	100	100	100	100	100	100	100	100	100	100	100	100	100
تفریحی_مبداء	Pearson Correlation	.682	.000	.782	.000	1	.000	.709	.000	.090	.000	-.344	.000	.160
	Sig. (2-tailed)	.000	1.000	.000	1.000		1.000	.000	1.000	.376	1.000	.000	1.000	.113
	N	100	100	100	100	100	100	100	100	100	100	100	100	100
تفریحی_مقصد	Pearson Correlation	.000	.682	.000	.782	.000	1	.000	.709	.000	.090	.000	-.344	.000
	Sig. (2-tailed)	1.000	.000	1.000	.000	1.000		1.000	.000	1.000	.376	1.000	.000	1.000
	N	100	100	100	100	100	100	100	100	100	100	100	100	100

۴-۳-۷ بررسی ضریب همبستگی در نرم افزار MATLAB

همچنین می توان همبستگی را با استفاده از کد معرفی شده در بخش ۴-۳-۸ در نرم افزار متلب ضرایب همبستگی را محاسبه نمود بگونه ای که در نهایت نرم افزار بر اساس ضریب همبستگی مورد نظر، میان پارامترهای همبسته پارامتری که واریانس داده های بالاتری دارد را انتخاب نموده و پارامتر های مناسب جهت تحلیل را در فایل اکسلی بعنوان خروجی ذخیره نماید.

در این تحقیق محاسبات با داده هایی که با یکدیگر ضریب همبستگی های بزرگتر مساوی با ۰/۵ و ۰/۶ و ۰/۷ و ۰/۸ و ۰/۹ داشتند پیش برده شد که پارامترهای مورد استفاده در هر حالت در بخش های ۵-۲-۱ و ۵-۲-۲ معرفی خواهند شد.

۴-۳-۸ کدهای استفاده شده

در این تحقیق از نرم افزار MATLAB R2013a استفاده شد. و کدهای دستوری استفاده شده

به شرح زیر می باشند:

- نرمال سازی داده ها

ابتدا باید فایل اکسلی که هر ستون مربوط به داده های یکی از پارامترها باشد را در پوشه ای که کد اسکریپت از آنجا فراخوانی شده است قرار داد. برای مثال در کد زیر، نام فایل اکسل اولیه ۸/۰.xlsx است. سپس نرم افزار متلب فایل اکسلی که داده های آن نرمال سازی شده اند را در همان پوشه با نام Normalization.xlsx ذخیره خواهد کرد

```
clc;clear;
Data=xlsread('8/0.xlsx');
[r,c]=size(Data);
for i=1:c
    MaxCol=max(Data(1:r,i));
    MinCol=min(Data(1:r,i));
    for j=1:r
        Data(j,i)=(Data(j,i)-MinCol)/(MaxCol-MinCol);
    end
end
xlswrite('Normalization.xlsx',Data);
```

- بررسی ضریب همبستگی داده ها

بوسیله این دستور، ابتدا فایل اکسلی که هر ستون حاوی داده های مربوط به یکی از متغیرها باشد و نام آن فایل Normalization.xlsx باشد، توسط نرم افزار خوانده می شود و در خط دهم مشخص می شود چه ضریب همبستگی ای مد نظر است که در اینجا ضریب همبستگی ۰.۸٪ مد نظر قرار داده شده و عدد ۰/۸ نوشته شده است. پس از آن نرم افزار متغیرهایی که بیش از ۰.۸٪ همبستگی داشته باشند را مشخص کرده و متغیری که واریانس داده های گسترده تری داشته باشد را انتخاب و مابقی را حذف می کند و در نهایت داده های مناسب برای ادامه تحلیل را در فایل

اکسلی به نام **FS_Data.xlsx** با چیدمان مشابه فایل ورودی (هر ستون مربوط به داده‌های یکی از متغیرها) ذخیره‌سازی خواهد کرد.

```
clc;clear;
Data=xlsread('Normalization.xlsx');
[r,c]=size(Data);
i=1;
while i<c-1
    j=i+1;
    while j<c
        rel1=corrcoef(Data(:,i),Data(:,j));
        mm=abs(rel1(1,2));
        if i~=j && mm>=8/0
            for aa=j+1:c
                Data(:,aa-1)=Data(:,aa);
            end
            Data(:,c)='';
            c=c-1;
        end
        j=j+1;
    end
    i=i+1;
end
xlswrite('FS_Data.xlsx',Data);
```

• جعبه ابزار شبکه عصبی متلب

با استفاده از دستور **nftool** در نرم افزار متلب می‌توان از جعبه ابزار شبکه عصبی متلب استفاده نمود. بصورت پیش فرض این شبکه عصبی دارای ۳ لایه می‌باشد که در لایه مخفی (میانی) دارای ۱۰ نورون جهت پردازش داده‌ها می‌باشد و داده‌ها در آن بصورت ۷۰٪ جهت آموزش و ۱۵٪ جهت اعتبار سنجی و ۱۵٪ جهت تست استفاده می‌شوند.

در این تحقیق از تنظیمات پیش فرض نرم افزار متلب استفاده شد بدین معنا که تغییری در تعداد نورون‌ها و درصدهای مربوط به تقسیم داده‌ها و الگوریتم آموزش (LM) و روش اعتبار

سنجی (MSE) ایجاد نشد.

- شبکه عصبی چندلایه بدون محدودیت داده اعتبارسنجی

این شبکه عصبی حاوی ۱۰ لایه می‌باشد و فایل اکسلی به نام 'FS_Data.xlsx' را بعنوان ورودی دریافت می‌کند که ستون‌ها از A متعلق به متغیرهای وابسته هستند و در نهایت در ستون پس از آخرین ستون متغیرهای وابسته، باید مقادیر مربوط به متغیر مستقل نوشته شود، برای مثال جدول ۴-۱۸ ستون‌های بدون هایلایت مربوط به داده‌های وابسته (Xiها) و ستون هایلایت شده مربوط به متغیر مستقل (Y) است.

در هر ردیف مقادیر به این صورت قرار می‌گیرند که اگر در معادله اولیه مان (معادله ۴-۲) مقادیر متغیرهای وابسته داخل یکی از ردیف‌های فایل اکسل را قرار دهیم، حاصل معادله (مقدار متغیر مستقل) برابر با مقدار مربوط به ستون متغیر مستقل در همان ردیف گردد.

$$aX_1 + bX_2 + \dots + cX_i = Y \quad \text{معادله ۴-۲}$$

جدول ۴-۱۵ بخشی از فایل ورودی به شبکه عصبی چندلایه بدون محدودیت داده اعتبارسنجی

A	B	BB	BC
9.36	9.36	2170	80043
9.36	7.47	2635	26457
9.36	29.34	2918	34120
9.36	16.95	3189	9205
9.36	607.57	5830	31884
9.36	465.5	2872	15579
9.36	19.21	6530	24955
9.36	20.65	7816	60083
9.36	2.11	1821	27972

در این شبکه ۸۰٪ تعداد داده‌ها بعنوان داده‌های آموزش و ۲۰٪ بعنوان داده‌های تست استفاده می‌شوند (باتوجه به اعداد ۰/۸ و ۰/۲ در خطوط ۸ و ۱۰) و از تابع train برای تعلیم شبکه

عصبی استفاده می‌گردد. و در نهایت نمودارهای مربوط به شبکه عصبی ترسیم شده و نتایج در فایل اکسلی با عنوان 'NNOut.xlsx' ذخیره می‌گردد که ستون آخر آن مربوط به داده های تخمین زده شده ($\hat{y}$) و ستون ماقبل آن مربوط به داده های تارگت (y) می‌باشند.

```
clc;clear;
layers=10;
Data=xlsread('FS_Data.xlsx');
[r,c]=size(Data);
Input = Data(:,1:c-1)';
Target = Data(:,c:c)';
net = newpr(Input,Target,10);
net.divideParam.trainratio=8/00;
net.divideParam.valratio=0;
net.divideParam.testratio=0.20;
[net,tr,Y,E,Pf,Af]= train(net,Input,Target);
W = sim(net,Input);
for i=1:c-1
    sum=0;
    for j=1:r
        a=W(j);
        b=Data(j,i);
        Data(j,i)=a*b;
    end
end
for i=1:r
    sum=0;
    for j=1:c-1
        sum=sum+Data(i,j);
    end
    Data(i,c+1)=sum;
end
xlswrite('NNOut.xlsx',Data);
```

- شبکه عصبی چندلایه دارای قابلیت حذف داده های حاوی خطا (داده های نویزی)

همانند شبکه عصبی چندلایه است که در بخش پیشین توضیح داده شد. با این تفاوت که در

بخش

```
for i=1:r
    if Data(i,c)<=5
        Result(cnt,:)=Data(i,:);
```

```

        cnt=cnt+1;
    end
end
xlswrite('Result.xlsx',Result);

```

عدد ۵ به معنای حذف داده هایی است که باعث خطای ۰.۵٪ و بیشتر می گردند. و با افزایش این مقدار، تعداد داده های کمتری حذف خواهند شد.

و در نهایت فایل اکسل 'Result.xlsx' نیز ذخیره می گردد که آخرین ستون آن میزان درصد نویز ای است که داده های آن ردیف ایجاد می کرده اند و ستون پیشین نیز تفاضل تارگت و مقدار تخمین زده شده $y^{\wedge}$ است و ستون پیشین مقدار $y^{\wedge}$ هستند.

```

clc;clear;
layers=10;
Data=xlsread('FS_Data.xlsx');
[r,c]=size(Data);
Input = Data(:,1:c-1)';
Target = Data(:,c:c)';

net = newpr(Input,Target,10);
net.divideParam.trainratio=8/00;
net.divideParam.valratio=0;
net.divideParam.testratio=0.20;
[net,tr,Y,E,Pf,Af]= train(net,Input,Target);
y=net(Input);
plot(Input,Target,'o',Input,y,'*');
Data(:,c+1)=y;
Data(:,c+2)=abs(Data(:,c)-Data(:,c+1));
Data(:,c+3)=(Data(:,c+2)./Data(:,c)).*100;
xlswrite('NNOOut.xlsx',Data);
[r,c]=size(Data);
cnt=1;
for i=1:r
    if Data(i,c)<=5
        Result(cnt,:)=Data(i,:);
        cnt=cnt+1;
    end
end
xlswrite('Result.xlsx',Result);

disp('End');

```

- شبکه عصبی پیش‌خور

در این شبکه عصبی ابتدا فایل اکسلی همانند کد مربوط به شبکه عصبی چندلایه وارد شبکه عصبی شده. این شبکه عصبی پیش‌خور حاوی دو لایه مخفی و در هر لایه ۵۰ نورون موجود می‌باشد. سپس ۷۰٪ داده‌ها برای آموزش، ۱۵٪ داده‌ها برای اعتبار سنجی و ۱۵٪ داده‌ها برای تست استفاده می‌شوند (مقادیر ۰/۷ و ۰,۱۵ در خطوط ۹ و ۱۰ و ۱۱). سپس نمودارهای مربوط به شبکه عصبی رسم خواهند شد و دو فایل اکسل ذخیره می‌گردد. فایل 'Y.xlsx' حاوی یک ستون است که نشان دهنده مقادیر تخمین زده شده توسط شبکه عصبی می‌باشد و فایل 'T.xlsx' حاوی یک ستون است که نشان دهنده مقادیر تارگت (متغیر مستقل) شبکه عصبی می‌باشد.

```
clc;
clear;
layers=[50 50];
input=xlsread('FS_Data.xlsx');
[r,c]=size(input);
datainput = input(:,1:c-1);
dataoutput = input(:,c:c);
net = newff(datainput,dataoutput,layers,{'logsig' 'logsig'});
net.divideParam.trainratio=7/0;
net.divideParam.valratio=0.15;
net.divideParam.testratio=0.15;

[net,tr,Y,E,Pf,Af]= train(net,datainput,dataoutput);
y = sim(net,datainput);

xlswrite('Y.xlsx',y);
xlswrite('T.xlsx',dataoutput);
disp('End');
```

۹-۳-۴ آماده سازی فایل اکسل ورودی

فایل اکسل استفاده شده در این تحقیق بگونه‌ای آماده سازی شده است تا هر ردیف نمایانگر مقادیر متغیرهای وابسته ای باشد که متناظر با مقدار عددی تعداد سفر(متغیر مستقل) بین دو ناحیه به ازای آن مقادیر متغیر وابسته باشد. به بیان ساده تر ابتدا جدول ۴-۱۹ ملاحظه شود. حال تعداد سفر از ناحیه ۱ به ناحیه ۱ و از ناحیه ۱ به ناحیه ۲&۱۱ (به اختصار ناحیه ۲ می‌نامیم) و از ناحیه ۲ به ناحیه ۱ بررسی میشود.

جدول ۴-۱۶ ماتریس سفر شهر اصفهان سال ۱۳۹۵

	1 2 & 11		3	4 5 & 13		6 7 & 14		8 & 12	9	10
1	80043	26457	34120	9205	31884	15579	24955	60083	27972	12286
2	27649	108740	10298	2526	10058	5992	7316	54154	13972	2745
3	33309	9994	128391	57810	32509	36283	66456	46817	7715	65048
4	9412	2570	59776	106888	18269	19752	11458	10189	1504	22946
5	33575	9534	35284	18490	447079	99441	21632	33799	9432	13521
6	16211	6071	37019	18917	99471	140845	18475	20739	4351	14511
7	26024	7390	69969	11695	23401	18468	237954	66206	2312	58204
8	62734	54545	50440	10889	35321	21479	65778	409513	8519	19368
9	28553	14140	7569	1564	9150	4099	2122	8330	46746	1694
10	12405	2659	68755	22983	13118	15124	55824	18912	1641	151574

در جدول ۴-۲۰ مشاهده می‌شود که متناظر با تعداد سفر ۴۳۰۸۰ مقادیر مربوط به پارامترهای مبداء مقادیر مربوط به ناحیه ۱ هستند و مقادیر مرتبط با پارامترهای مقصد مقادیر مربوط به ناحیه ۱ هستند اما متناظر با تعداد سفر ۷۴۴۵۷ که تعداد سفر ناحیه ۱ به ناحیه ۲ است، مقادیر مربوط به پارامترهای مبداء مقادیر مربوط به ناحیه ۱ هستند و مقادیر مرتبط با پارامترهای مقصد مقادیر مربوط به ناحیه ۲ هستند و مابقی داده‌ها نیز به همین منوال تعیین شده اند. لذا می‌توان از این فایل اکسل در تخمین استفاده کرد

جدول ۴-۱۷ بخشی از ماتریس ورودی به نرم‌افزار متلب جهت فرآیند تخمین

AY	AZ	BA	BB	BC
وضع فعالیت کد ۷ (مبدأ)	وضع فعالیت کد ۷ (مقصد)	وضع فعالیت کد ۸ (مبدأ)	وضع فعالیت کد ۸ (مقصد)	ماتریس سفر
7537	7537	2170	2170	80043
7537	7269	2170	2635	26457
7537	10752	2170	2918	34120
7537	12419	2170	3189	9205
7537	28321	2170	5830	31884
7537	13408	2170	2872	15579
7537	18583	2170	6530	24955
7537	27122	2170	7816	60083
7537	4858	2170	1821	27972
7537	14458	2170	4401	12286
7269	7537	2635	2170	27649
7269	7269	2635	2635	108740
7269	10752	2635	2918	10298
7269	12419	2635	3189	2526
7269	28321	2635	5830	10058
7269	13408	2635	2872	5992
7269	18583	2635	6530	7316

در این فصل به بررسی روش تحقیق پرداخته شد و نتایج در فصل پنجم معرفی و بررسی

خواهند شد.

فصل ۵: مدلسازی و نتایج تحقیق

۵-۱ مقدمه

روش‌های مورد بررسی در این فصل عبارتند از:

- (۱) شبکه عصبی جعبه ابزار متلب (۲) شبکه عصبی چندلایه بدون داده‌های اعتبار سنجی (۳) شبکه عصبی با حذف داده‌های نویزی (۴) شبکه عصبی پیشخور
- پارامترها با توجه به ضریب همبستگی میانشان تفکیک شدند و فرآیند تخمین بر روی دسته پارامترهای با ضریب همبستگی کمتر از ۰/۵ و ۰/۶ و ۰/۷ و ۰/۸ و ۰/۹ انجام شد و بهترین نتایج موفقیت آمیز در ادامه مورد بررسی و تحلیل قرار گرفتند.

۵-۲ تحلیل نمودارهای خروجی

در ابتدا شبکه عصبی داده‌ها را به سه دسته بصورت تصادفی و با توجه به مقداری که در ابتدا برایش تعریف می‌شود، تقسیم بندی می‌کند. بصورت پیشفرض برای تقسیم داده‌ها، ۷۰٪ از کل تعداد داده‌ها جهت فرآیند آموزش، ۱۵٪ جهت فرآیند اعتبار سنجی و ۱۵٪ جهت فرآیند تست استفاده می‌شوند و در این تحقیق نیز در دو حالت داده‌ها تقسیم بندی شدند. هنگامی که از اعتبار سنجی استفاده شد، تقسیم داده‌ها بصورت ۷۰٪ جهت آموزش، ۱۵٪ جهت اعتبار سنجی و ۱۵٪ جهت تست بود و هنگامی که اعتبار سنجی حین فرآیند آموزش انجام نشد، میزان ۸۰٪ از داده‌ها جهت فرآیند آموزش و ۲۰٪ جهت فرآیند تست انتخاب شدند.

در نمودارهای خروجی نرم افزار MATLAB 2013 محور افقی نشان دهنده Epoch است. یک دوره (Epoch) به یک تکثیر رفت و برگشت در شبکه اشاره دارد. یعنی یک دوره برابر با یک رفت و برگشت ورودی در کل شبکه است. تعداد دوره‌هایی که برای یادگیری شبکه استفاده می‌کنید کاملاً دست خودتان است. درست است که داشتن دوره‌های بیشتر منجر به دقت بیشتر در شبکه می‌شود، ولی زمان یادگیری شبکه را هم افزایش می‌دهد. به علاوه، باید توجه داشته باشید که اگر

تعداد دوره‌ها بیش از حد زیاد باشد، ممکن است هضم آن برای شبکه سنگین باشد، یا اصطلاحاً شبکه «over-fit» شود.

محور قائم در این نمودارها نشان دهنده مقدار MSE است (معادله ۵-۱). خطای میانگین مربعات (Mean squared error) روشی برای برآورد میزان خطاست که در واقع تفاوت بین مقادیر تخمینی و آنچه تخمین زده شده است.

$$MSE = \frac{1}{n} \sum_{i=1}^n (\bar{Y}_i - Y_i)^2 \quad \text{معادله ۵-۱}$$

n = تعداد داده‌ها ، Y = مقادیر واقعی متغیر مستقل $\bar{Y}$ = مقادیر تخمین زده شده متغیر مستقل
نمودار با سه خط ترسیم شده است که خط آبی رنگ نشانگر میزان MSE برای داده‌های آموزش در هر دوره است و به همین ترتیب، رنگ سبز نشانگر MSE داده‌های اعتبار سنجی و رنگ قرمز نشانگر MSE داده‌های تست می‌باشند.

در صورتی که روند خطوط آبی، سبز و قرمز بصورت خطی نسبتاً افقی درآمده باشد می‌توان گفت این خطوط همگرا شده‌اند و می‌توان به نتیجه شبکه عصبی اعتماد نمود. و در صورتی که همگرایی رخ ندهد، از آنجا که مشخص نیست در دوره‌های بعدی، شبکه عصبی ممکن بوده است چه رفتاری از خود نشان دهد و شاید در دوره‌های بعدی خطا افزایش می‌یافت، لذا به نتیجه بدست آمده نمی‌توان اعتماد کرد و باید با حذف محدودیت و افزایش تعداد دوره‌های مجاز برای شبکه عصبی، رفتار آن را در ادامه فرآیند نیز مورد بررسی قرار داد. بدین منظور روش‌های بدون محدودیت داده اعتبار سنجی نیز در این تحقیق بررسی شدند.

۵-۲-۱ داده های شهر مشهد

ابتدا با استفاده از نرم افزار متلب داده های دارای ضریب همبستگی ۰/۵ و ۰/۶ و ۰/۷ و ۰/۸ و ۰/۹ را تعیین شدند و قابل ذکر است که نرمال سازی تاثیری بر پارامتر انتخاب شده توسط نرم افزار متلب نداشت و در هر دو حالت استفاده از داده های نرمال سازی شده و داده های نرمال سازی نشده، خروجی نرم افزار پارامترهای یکسانی بودند.

پس از بررسی ضرایب همبستگی، در ادامه پارامترها معرفی می شوند تا مشخص شود که با در نظر گرفتن مقادیر مختلف ضرایب همبستگی، از کدام پارامترها در تخمین استفاده شده است.

• داده های بررسی شده با در نظر گرفتن ضریب همبستگی ۰/۵

اداری (مبداء)، اداری (مقصد)، تفریحی (مقصد)، حمل و نقل (مقصد)، تجاری (مبداء)، صنعتی (مبداء)، جمعیت کل (مبداء)، جمعیت کل (مقصد)، جمعیت زن (مقصد)، جمعیت ۶ ساله و بیشتری سواد (مقصد)، جمعیت بیکار زن (مقصد)، وضع فعالیت کد ۸ (سایر) (مقصد)

• داده های بررسی شده با در نظر گرفتن ضریب همبستگی ۰/۶

اداری (مبداء)، اداری (مقصد)، تفریحی (مقصد)، درمانی (مبداء)، درمانی (مقصد)، حمل و نقل (مبداء)، حمل و نقل (مقصد)، تجاری (مبداء)، تجاری (مقصد)، جمعیت کل (مبداء)، جمعیت کل (مقصد)، جمعیت زن (مقصد)، جمعیت ۶ ساله و بیشتری سواد (مقصد)، جمعیت ۱۰ ساله و بیشتر زن (مقصد)، وضع فعالیت کد ۸ (سایر) (مقصد)

• داده های بررسی شده با در نظر گرفتن ضریب همبستگی ۰/۷

اداری (مبداء)، اداری (مقصد)، تفریحی (مبداء)، تفریحی (مقصد)، درمانی (مبداء)، درمانی (مقصد)، حمل و نقل (مبداء)، حمل و نقل (مقصد)، تجاری (مبداء)، تجاری (مقصد)، مسکونی (مبداء)، مسکونی (مقصد)، جمعیت کل (مبداء)، جمعیت کل (مقصد)، جمعیت زن (مقصد)، جمعیت ۶ ساله و بیشتری سواد (مبداء)، جمعیت ۶ ساله و بیشتری سواد (مقصد)، جمعیت ۱۰ ساله و

بیشترزن(مقصد)، وضع فعالیت کد ۸ (سایر)(مقصد)

• داده های بررسی شده با در نظر گرفتن ضریب همبستگی ۰/۸

اداری(مبداء)، اداری(مقصد)، آموزشی(مبداء)، آموزشی(مقصد)، درمانی(مبداء)، درمانی(مقصد)، حمل و نقل(مبداء)، حمل و نقل(مقصد)، تجاری(مبداء)، تجاری(مقصد)، جمعیت کل(مبداء)، جمعیت کل(مقصد)، جمعیت زن(مقصد)، جمعیت ۶ ساله و بیشتر بی سواد(مبداء)، جمعیت ۶ ساله و بیشتر بی سواد(مقصد)، جمعیت ۱۰ ساله و بیشترزن(مقصد)، جمعیت شاغل زن(مبداء)، جمعیت بیکار زن(مقصد)، وضع فعالیت کد ۸ (سایر)(مقصد)

• داده های بررسی شده با در نظر گرفتن ضریب همبستگی ۰/۹

اداری(مبداء)، اداری(مقصد)، آموزشی(مبداء)، آموزشی(مقصد)، تفریحی(مبداء)، تفریحی(مقصد)، درمانی(مبداء)، درمانی(مقصد)، حمل و نقل(مبداء)، حمل و نقل(مقصد)، تجاری(مبداء)، تجاری(مقصد)، مسکونی(مبداء)، مسکونی(مقصد)، جمعیت کل(مبداء)، جمعیت کل(مقصد)، جمعیت زن(مقصد)، جمعیت ۶ ساله و بیشتر بی سواد(مبداء)، جمعیت ۶ ساله و بیشتر بی سواد(مقصد)، جمعیت ۱۰ ساله و بیشترزن(مقصد)، جمعیت شاغل زن(مبداء)، جمعیت شاغل زن(مقصد)، وضع فعالیت کد ۷ (دارای درآمد بدون کار)(مقصد)، وضع فعالیت کد ۸ (سایر)(مقصد).

در ادامه به بررسی شیوه تخمین با روش های مورد نظر در این تحقیق و همچنین تحلیل بهترین نتایج بدست آمده در هر روش پرداخته شده است.

۵-۲-۱- شبکه عصبی جعبه ابزار متلب

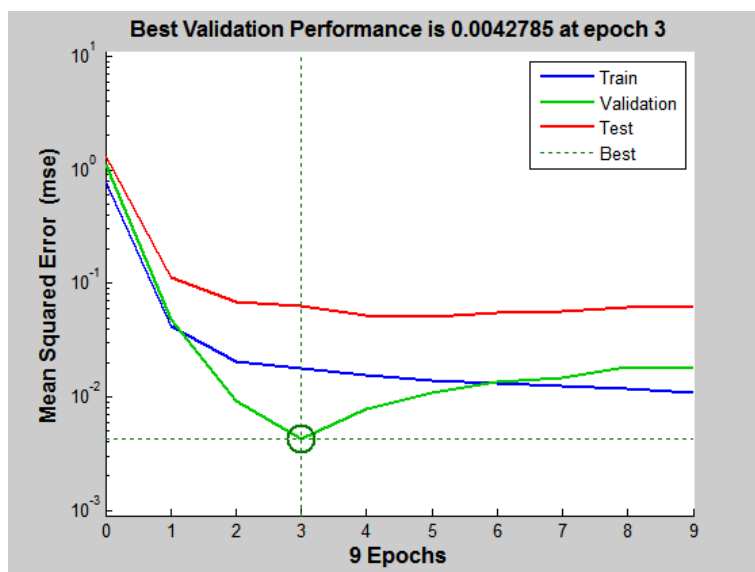
قابل ذکر است که برای ضرایب همبستگی ۰/۵ تا ۰/۹ بر روی دو حالت داده های نرمال سازی شده و نرمال سازی نشده تخمین ماتریس سفر انجام شد که در ادامه این فصل تنها به تحلیل و

بررسی مناسب‌ترین نتایج قابل قبول اشاره می‌شود و ۷۳ مورد مابقی در فایل موجود در CD ضمیمه با فرمت docx. به نام ((ضمیمه)) آورده شده است.

جهت شروع فرآیند تخمین ابتدا داده‌های شهر مشهد نرمال‌سازی شدند و سپس داده‌هایی که ضریب همبستگی بیشتر از ۰/۹ داشتند شناسایی شدند و آن پارامتری که واریانس داده‌های بیشتری داشت انتخاب شد (پارامترهای تعیین شده در بخش ۵-۱-۲). سپس در نرم افزار متلب با استفاده از دستور nftool جعبه ابزار شبکه عصبی متلب فراخوانی شد و پس از آن داده‌ها بدون ایجاد تغییر در تنظیمات پیش فرض جعبه ابزار متلب، وارد شبکه عصبی پیش فرض شدند و فرآیند تخمین بر روی آنها انجام شد.

با توجه به شکل ۵-۱، پس از سه دوره با توجه به محدودیت داده‌های اعتبارسنجی شبکه عصبی متوقف شد و بهترین MSE گزارش شده برابر با ۰/۰۰۴۲۷۸۵ و در نتیجه RMSE برابر با ۰,۰۶۵۴۱۰۲۴۴ بود (RMSE برابر با ریشه دوم MSE است) که مقدار بسیار مناسبی است. همچنین با توجه به نمودار از آنجایی که خطوط بنظر در دامنه کمی در نوسان اند و می توان گفت به مقدار عددی خاصی همگرا شده اند، می‌توان به نتایج شبکه عصبی اعتماد کرد.

در بخش ۵-۲-۱-۲ در روش شبکه عصبی بدون محدودیت داده های اعتبارسنجی، این موضوع بررسی شده است که اگر شبکه عصبی متوقف نشده بود و ۱۰۰۰ دوره تلاش انجام می‌شد آیا همچنان خطوط MSE داده های آموزش و تست به عدد خاصی همگرا خواهند بود یا خیر و نتیجه حاصل شد که آری این خطوط بصورت همگرا ادامه مسیر خواهند داد و نتایج بدست آمده کاملاً قابل اعتماد می‌باشند.



شکل ۵-۱ نتایج تخمین ماتریس سفر شهر مشهد با استفاده از شبکه عصبی پیش فرض جعبه ابزار متلب بوسیله داده های با ضریب همبستگی کمتر از ۰/۹

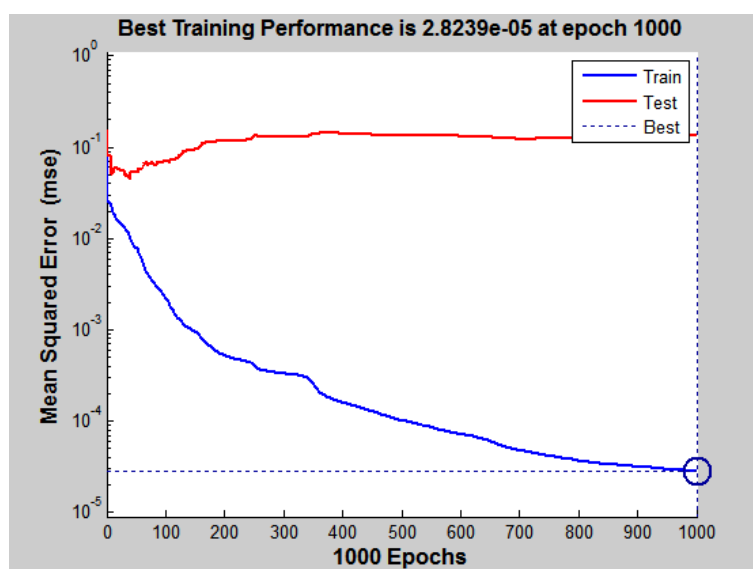
سپس فرآیند تخمین بر روی داده های نرمال سازی نشده انجام شد و سپس داده هایی که ضریب همبستگی بیشتر از ۰/۸ داشتند شناسایی شدند و آن پارامتری که واریانس داده های بیشتری داشت انتخاب شد (پارامترهای تعیین شده در بخش ۵-۱-۲). سپس در نرم افزار متلب با استفاده از دستور nftool جعبه ابزار شبکه عصبی متلب فراخوانی شد و پس از آن داده ها بدون ایجاد تغییر در تنظیمات پیش فرض جعبه ابزار متلب، وارد شبکه عصبی پیش فرض شدند و فرآیند تخمین بر روی آنها انجام شد.

۵-۲-۱-۲ شبکه عصبی چندلایه بدون داده های اعتبار سنجی

ابتدا داده ها شهر مشهد نرمال سازی شد و سپس داده هایی که ضریب همبستگی بیشتر از ۰/۹ داشتند شناسایی شده و آن پارامتری که واریانس داده های بیشتری داشت انتخاب شد (پارامترهای تعیین شده در بخش ۵-۱-۲). سپس در نرم افزار متلب با استفاده از بخش script معماری شبکه عصبی طراحی شده که در بخش ۴-۳-۹ معرفی شد در نرم افزار متلب بارگذاری

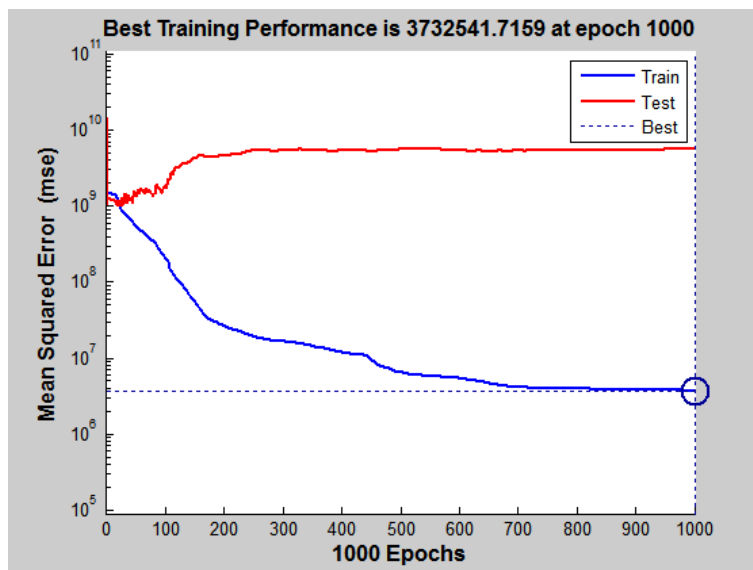
شد و سپس داده ها وارد شبکه شدند و تخمین ماتریس سفر انجام شد.

با توجه به شکل ۵-۳، پس از ۱۰۰۰ دوره با توجه به محدودیت تعداد دوره، شبکه عصبی متوقف شد و بهترین MSE گزارش شده برابر با 0.000028239 بود که مقدار بسیار مناسبی است. همچنین با توجه به نمودار از آنجایی که خطوط در ادامه مسیر شان در دامنه کمی در نوسان اند و می توان گفت به مقدار عددی خاصی همگرا شده اند، می توان به نتایج شبکه عصبی اعتماد کرد.



شکل ۵-۲ نتایج تخمین ماتریس سفر شهر مشهد با استفاده از شبکه عصبی طراحی شده بدون محدودیت داده های اعتبارسنجی بوسیله داده های با ضریب همبستگی کمتر از 0.9

با توجه به شکل ۵-۳، پس از ۱۰۰۰ دوره با توجه به محدودیت تعداد دوره، شبکه عصبی متوقف شد و MSE این تخمین با توجه به بزرگ بودن مقادیر عددی داده ها موجه بنظر می رسد و خطوط بنظر در دامنه کمی در نوسان اند و می توان گفت به مقدار عددی خاصی همگرا شده اند، لذا می توان به نتایج شبکه عصبی اعتماد کرد.



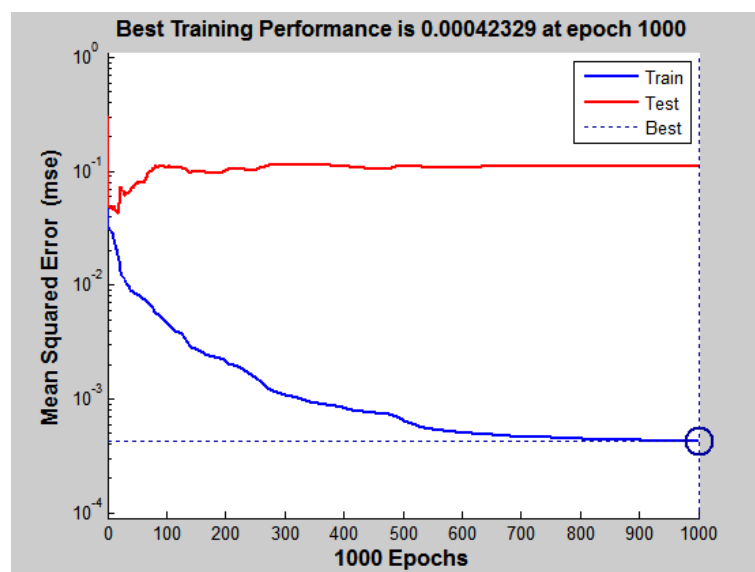
شکل ۳-۵ شبکه عصبی چندلایه بدون داده های اعتبار سنجی (داده های بدون نرمال سازی) داده های با ضریب همبستگی ۰/۷

۳-۱-۲-۵ شبکه عصبی با حذف داده های نویزی

در این روش ابتدا باید درصد نویز مورد نظر را تعیین نمود تا نرم افزار داده هایی که خطای بالاتری دارند را حذف کند و تخمین اصلی را با داده هایی که درصد نویز کمتری از مقدار مورد نظر ایجاد کنند ادامه دهد. لذا ابتدا درصد نویز را ۵۰٪ تعیین می کنیم سپس داده ها شهر مشهد نرمال سازی شده و داده هایی که ضریب همبستگی بیشتر از ۰/۶ داشتند شناسایی شده و آن پارامتری که واریانس داده های بیشتری داشت انتخاب شد (پارامترهای تعیین شده در بخش ۵-۱-۲). سپس در نرم افزار متلب با استفاده از بخش script معماری شبکه عصبی طراحی شده که در بخش ۴-۳-۹ معرفی شد بارگذاری شد و داده هایی که بعنوان داده های ایجاد کننده نویز ۵۰٪ شناخته می شدند مشخص شده و داده های مناسب، وارد شبکه شدند و تخمین ماتریس سفر انجام شد.

در فایل خروجی این شبکه عصبی، داده های مورد بررسی (در بخش ۴-۳-۹ نحوه تفکیک داده

های نویزی از مجموع داده‌ها توضیح داده شد) تعیین می‌شوند. هر ردیف در فایل اکسل ورودی به منزله یک معادله است که مقادیر متغیرهای مستقل متناظر با مقدار متغیر مستقل موجود در ستون آخر را در خود جای داده است. با بررسی فایل اکسل خروجی شبکه عصبی می‌توان فهمید که از مجموع ۱۴۴ معادله اولیه تعداد معادلاتی که بعنوان نویز تشخیص داده شده و پیش از فرآیند تخمین حذف شده بودند تعداد ۹۳ معادله بوده اند و تعداد ۵۱ معادله بررسی شده بوده اند. با توجه به شکل ۴-۵ پس از ۱۰۰۰ دوره با توجه به محدودیت تعداد دوره، شبکه عصبی متوقف شد و بهترین MSE گزارش شده برابر با ۰/۰۰۰۴۲۳۲۹ بود که مقدار مناسبی است. همچنین با توجه به نمودار از آنجایی که خطوط در ادامه مسیر شان در دامنه کمی در نوسان اند و می‌توان گفت به مقدار عددی خاصی همگرا شده اند، می‌توان به نتایج شبکه عصبی اعتماد کرد.



شکل ۴-۵ نتایج تخمین ماتریس سفر شهر مشهد با استفاده از شبکه عصبی طراحی شده دارای قابلیت حذف

داده‌های نویزی بوسیله داده‌های با ضریب همبستگی کمتر از ۰/۶

۵-۲-۲ داده‌های شهر اصفهان

ابتدا با استفاده از نرم افزار متلب داده‌های دارای ضریب همبستگی ۰/۵ و ۰/۶ و ۰/۷ و ۰/۸ و

۰/۹ را تعیین شدند:

• داده های بررسی شده با در نظر گرفتن ضریب همبستگی ۰/۵

اداری(مبداء)، اداری(مقصد)، تفریحی(مقصد)، حمل و نقل(مبداء)، حمل و نقل(مقصد)،
تجاری(مبداء)، تجاری(مقصد)، مسکونی(مقصد)، تعداد کل خانوار(مبداء)، جمعیت شاغل
مرد(مقصد)، جمعیت بیکار مرد(مبداء)، وضع فعالیت کد ۸ (سایر)(مبداء)، وضع فعالیت کد ۸
(سایر)(مقصد)

• داده های بررسی شده با در نظر گرفتن ضریب همبستگی ۰/۶

اداری(مبداء)، اداری(مقصد)، تفریحی(مقصد)، حمل و نقل(مبداء)، حمل و نقل(مقصد)،
تجاری(مبداء)، تجاری(مقصد)، مسکونی(مقصد)، تعداد کل خانوار(مبداء)، جمعیت ۶ ساله و
بیشتری سواد(مبداء)، جمعیت شاغل زن(مقصد)، وضع فعالیت کد ۸ (سایر)(مبداء)، وضع فعالیت
کد ۸ (سایر)(مقصد)

• داده های بررسی شده با در نظر گرفتن ضریب همبستگی ۰/۷

اداری(مبداء)، اداری(مقصد)، تفریحی(مبداء)، تفریحی(مقصد)، حمل و نقل(مبداء)، حمل و
نقل(مقصد)، تجاری(مبداء)، تجاری(مقصد)، مسکونی(مقصد)، تعداد کل خانوار(مبداء)، جمعیت ۶
ساله و بیشتری سواد(مبداء)، جمعیت ۱۰ ساله و بیشتر زن(مقصد)، جمعیت شاغل زن(مبداء)، وضع
فعالیت کد ۷ (دارای درآمد بدون کار)(مبداء)، وضع فعالیت کد ۸ (سایر)(مقصد)

• داده های بررسی شده با در نظر گرفتن ضریب همبستگی ۰/۸

اداری(مبداء)، اداری(مقصد)، تفریحی(مبداء)، تفریحی(مقصد)، درمانی(مبداء)، درمانی(مقصد)،
حمل و نقل(مبداء)، حمل و نقل(مقصد)، تجاری(مبداء)، تجاری(مقصد)، صنعتی(مبداء)،
صنعتی(مقصد)، جمعیت کل(مبداء)، جمعیت کل(مقصد)، جمعیت زن(مقصد)، جمعیت ۶ ساله و

بیشتری سواد(مقصد)، جمعیت بیکار زن(مقصد)، وضع فعالیت کد ۸ (سایر)(مقصد)

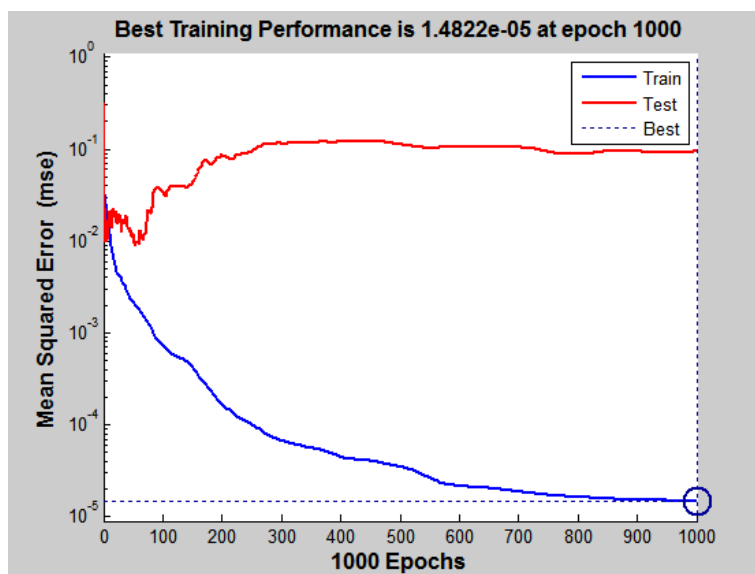
• داده های بررسی شده با در نظر گرفتن ضریب همبستگی ۰/۹

اداری(مبداء)، اداری(مقصد)، آموزشی(مبداء)، آموزشی(مقصد)، تفریحی(مبداء)، تفریحی(مقصد)، حمل و نقل(مبداء)، حمل و نقل(مقصد)، تجاری(مبداء)، تجاری(مقصد)، صنعتی(مبداء)، صنعتی(مقصد)، مسکونی(مبداء)، مسکونی(مقصد)، جمعیت کل(مبداء)، جمعیت کل(مقصد)، جمعیت زن(مقصد)، جمعیت ۶ ساله و بیشتری سواد(مبداء)، جمعیت ۶ ساله و بیشتری سواد(مقصد)، جمعیت شاغل زن(مبداء)، جمعیت شاغل زن(مقصد)، وضع فعالیت کد ۷ (دارای درآمد بدون کار)(مقصد)، وضع فعالیت کد ۸ (سایر)(مقصد)

۵-۲-۱ شبکه عصبی چندلایه بدون داده های اعتبار سنجی

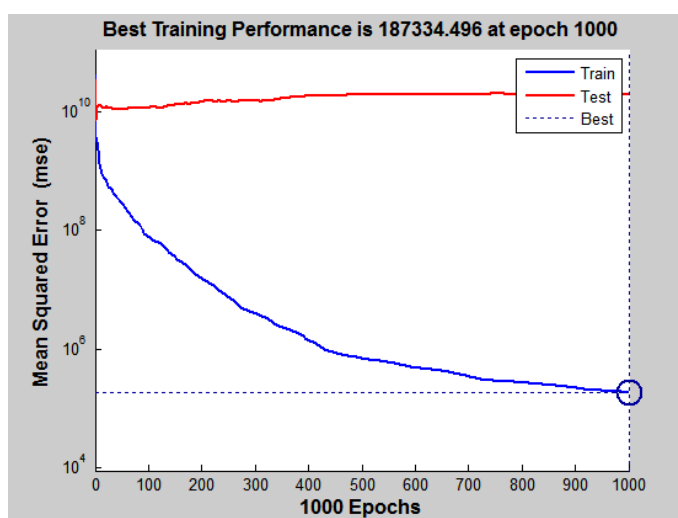
ابتدا داده ها شهر اصفهان نرمال سازی شد و سپس داده هایی که ضریب همبستگی بیشتر از ۰/۸ داشتند شناسایی شده و آن پارامتری که واریانس داده های بیشتری داشت انتخاب شد(پارامترهای تعیین شده در بخش ۵-۱-۳). سپس در نرم افزار متلب با استفاده از بخش script معماری شبکه عصبی طراحی شده که در بخش ۴-۳-۹ معرفی شد در نرم افزار متلب بارگذاری شد و سپس داده ها وارد شبکه شدند و تخمین ماتریس سفر انجام شد.

با توجه به شکل ۵-۵، پس از ۱۰۰۰ دوره با توجه به محدودیت تعداد دوره، شبکه عصبی متوقف شد و بهترین MSE گزارش شده برابر با ۰/۰۰۰۰۱۴۸۲۲ بود که مقدار مناسبی است. همچنین با توجه به نمودار از آنجایی که خطوط در ادامه مسیر شان در دامنه کمی در نوسان اند و می توان گفت به مقدار عددی خاصی همگرا شده اند، می توان به نتایج شبکه عصبی اعتماد کرد



شکل ۵-۵ نتایج تخمین ماتریس سفر شهر اصفهان با استفاده از شبکه عصبی طراحی شده بدون محدودیت داده های اعتبارسنجی بوسیله داده های با ضریب همبستگی کمتر از ۰/۸

با توجه به شکل ۵-۶، پس از ۱۰۰۰ دوره با توجه به محدودیت تعداد دوره، شبکه عصبی متوقف شد و MSE این تخمین با توجه به بزرگ بودن مقادیر عددی داده ها موجه بنظر می رسد و خطوط بنظر در دامنه کمی در نوسان اند و می توان گفت به مقدار عددی خاصی همگرا شده اند، لذا می توان به نتایج شبکه عصبی اعتماد کرد. MSE این نتیجه برابر با $187334.496/496$ بود.



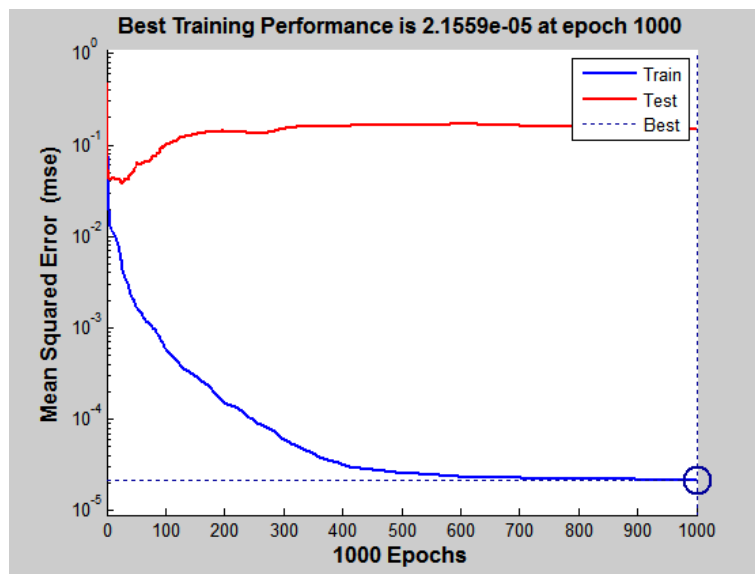
شکل ۵-۶ شبکه عصبی چندلایه بدون داده های اعتبارسنجی (داده های بدون نرمال سازی) داده های با ضریب همبستگی ۰,۹

۵-۲-۲-۲ شبکه عصبی با حذف داده های نویزی

ابتدا داده ها شهر اصفهان نرمال سازی شد و سپس داده هایی که ضریب همبستگی بیشتر از ۰/۸ داشتند شناسایی شده و آن پارامتری که واریانس داده های بیشتری داشت انتخاب شد (پارامترهای تعیین شده در بخش ۵-۱-۳). سپس در نرم افزار متلب با استفاده از بخش script معماری شبکه عصبی طراحی شده که در بخش ۴-۳-۹ معرفی شد در نرم افزار متلب بازگذاری شد و سپس داده هایی که بعنوان داده های ایجاد کننده نویز ۵۰٪ شناخته می شدند مشخص شده و داده های مناسب، وارد شبکه شدند و تخمین ماتریس سفر انجام شد.

در فایل خروجی این شبکه عصبی، داده های مورد بررسی (در بخش ۴-۳-۹ نحوه تفکیک داده های نویزی از مجموع داده ها توضیح داده شد) تعیین می شوند و با بررسی فایل اکسل خروجی شبکه عصبی می توان فهمید که تعداد معادلاتی که بعنوان نویز تشخیص داده شده و پیش از فرآیند تخمین حذف شده بودند تعداد ۳۷ معادله بوده اند و تعداد معادلات بررسی شده ۶۳ بوده اند.

با توجه به شکل ۵-۷، پس از ۱۰۰۰ دوره با توجه به محدودیت تعداد دوره، شبکه عصبی متوقف شد و بهترین MSE گزارش شده برابر با ۰/۰۰۰۰۲۱۵۵۹ بود که مقدار مناسبی است. همچنین با توجه به نمودار از آنجایی که خطوط در ادامه مسیر شان در دامنه کمی در نوسان اند و می توان گفت به مقدار عددی خاصی همگرا شده اند، می توان به نتایج شبکه عصبی اعتماد کرد

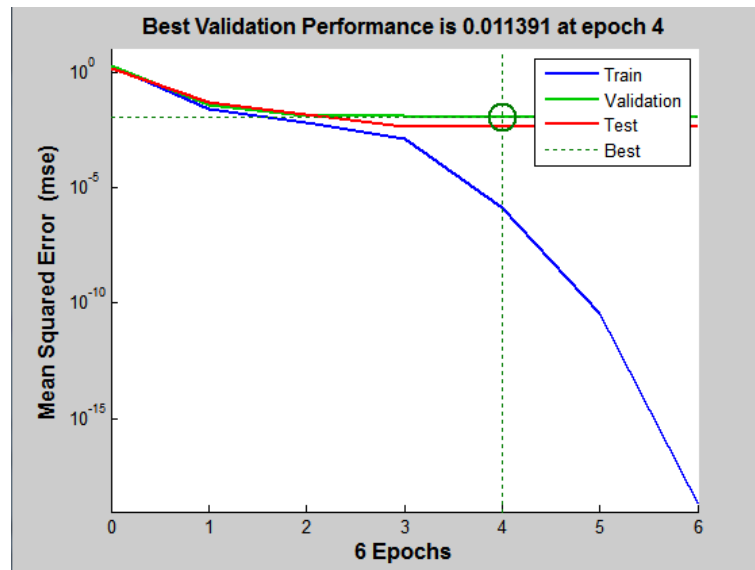


شکل ۵-۷ نتایج تخمین ماتریس سفر شهر اصفهان با استفاده از شبکه عصبی طراحی شده دارای قابلیت حذف داده‌های نویزی بوسیله داده‌های با ضریب همبستگی کمتر از ۰/۸

۳-۲-۲-۵ شبکه عصبی پیشخور

ابتدا داده‌ها شهر اصفهان نرمال سازی شد و سپس داده‌هایی که ضریب همبستگی بیشتر از ۰/۹ داشتند شناسایی شده و آن پارامتری که واریانس داده‌های بیشتری داشت انتخاب شد (پارامترهای تعیین شده در بخش ۵-۱-۳). سپس در نرم افزار متلب با استفاده از بخش script معماری شبکه عصبی طراحی شده که در بخش ۴-۳-۹ معرفی شد در نرم افزار متلب بازگذاری شد و سپس داده‌ها وارد شبکه شدند و تخمین ماتریس سفر انجام شد.

با توجه به شکل ۵-۸، پس از ۶ دوره با توجه به محدودیت تعداد دوره، شبکه عصبی متوقف شد و بهترین MSE گزارش شده برابر با ۰/۰۱۱۳۹۱ بود که مقدار مناسبی است. همچنین با توجه به نمودار از آنجایی که خطوط در ادامه مسیر شان در دامنه کمی در نوسان اند (خط آبی با توجه به نوع شبکه عصبی به صفر میل خواهد کرد) و می توان گفت به مقدار عددی خاصی همگرا شده اند، می توان به نتایج شبکه عصبی اعتماد کرد.



شکل ۵-۸ نتایج تخمین ماتریس سفر شهر اصفهان با استفاده از شبکه عصبی پیشخور بوسیله داده های با ضریب همبستگی کمتر از ۰/۹

۵-۳ مقایسه ماتریس سفر تخمین زده شده و اولیه

در این بخش هدف آن است تا بهترین ماتریس سفر بدست آمده توسط نتایج پیشین را با ماتریس اولیه آن مقایسه کرد. هنگامی که در شبکه عصبی از داده های نرمالسازی شده استفاده شود صرفاً می توان دقت روش ها را مقایسه کرد و از آنجا که باید اعداد متناظر با x ها و y بصورت نرمال شده در بیایند، و از آنجا که نمی توان مجدد این مقادیر را به مقادیر نرمالسازی نشده بازگردانی کرد لذا ماتریس سفرهای پیش رو بر اساس مقادیر نرمالسازی نشده تخمین زده شدند. بر اساس نتایج پیشین می توان بیان داشت که روش شبکه عصبی بدون محدودیت داده اعتبارسنجی نتایج بسیار مناسب تری نسبت به باقی روش ها ارائه می دهد.

۵-۳-۱ ماتریس سفر تخمین زده شده شهر مشهد

در این بخش ماتریس سفر اولیه شهر مشهد مربوط به سال ۱۳۹۵ ارائه و سپس ماتریس سفر تخمین زده شده ارائه خواهند شد. ابتدا در جدول ۵-۱ ماتریس سفر سال ۱۳۹۵ شهر مشهد ارائه

شده است و T مففی از کلمه Target به معنای هدف در فرآیند تخمین با شبکه عصبی است.

جدول ۵-۱ ماتریس سفر Target شهر مشهد

T	1	2	3	4	5	6	7	8	9	10	11	12
1	1113 37.9	58343 .69	18357 .45	3769. 531	2457. 581	4593. 683	14702 .05	29058 .39	27857 .52	13761 .38	19834 .57	4745. 723
2	5544 5.06	14959 2.8	48432 .51	9065. 397	3837. 867	4695. 815	8867. 236	18692 .46	21630 .47	33106 .38	26661 .3	3186. 527
3	2031 6.09	49557 .6	13978 9.3	44341 .41	10004 .78	11936 .43	12360 .8	14950 .45	9188. 542	11825 .7	9115. 957	1338. 852
4	4367 .845	10394 .28	44688 .39	92851 .65	14909 .18	14376 .82	9799. 127	5883. 19	2189. 383	2524. 265	1852. 237	274.4 003
5	2532 .553	4347. 432	10443 .24	15004 .91	21256 .1	12248 .31	7600. 979	3563. 435	1406. 24	1056. 653	914.3 023	118.8 098
6	5729 .416	5973. 76	13152 .72	14570 .67	12653 .49	65638 .62	24634 .55	11353 .26	3977. 476	1785. 101	19766 6	205.0 642
7	1588 2	10514 .2	12838 .29	10118 .86	8007. 324	22488 .45	10692 3.4	33297 .89	10603 .9	4439. 288	6878. 525	615.5 871
8	3151 6.18	19942 .08	18488 .81	7228. 695	4041. 517	9410. 718	33052 .56	55222 .41	16026 .2	8573. 058	7395. 27	1099. 813
9	3804 3.16	22767 .87	10241 .77	2553. 403	2320. 248	4445. 271	15401 .5	23651 .7	15409 6.7	20226 .33	49086 .52	6405. 873
10	1740 7.57	34695 .78	13981 .54	2855. 835	1081. 416	1618. 986	3945. 32	8798. 3	19817 .02	15672 1.9	35327 .91	12868 .93
11	2342 8.75	24283 .83	8369. 818	1761. 822	754.0 39	1520. 389	5172. 285	10203 .26	40936 .62	27672 .26	88861 .41	4237. 591
12	1994 .888	2865. 429	1429. 36	269.4 181	115.5 312	182.7 41	572.1 462	1058. 389	4239. 863	9531. 414	4221. 25	6637. 372

سپس در جدول ۵-۲ ماتریس سفر تخمین زده شده مشاهده می‌گردد. در این ماتریس O مخفف

کلمه OutPuts به معنای خروجی شبکه عصبی می‌باشد.

جدول ۵-۲ ماتریس سفر OutPut شهر مشهد

O	1	2	3	4	5	6	7	8	9	10	11	12
1	1113 33.5	5834 3.56	6314 .484	6314 .486	2455 .39	4594 .638	4002 3.92	2906 5.7	2814 1.69	3474 2.22	6656 8.15	4744 .705
2	3475 7.84	1495 92.8	4843 2.47	6314 .438	5318 .961	5635 .646	8868 .706	1745 8.12	1960 0.07	3474 2.24	2666 2.35	5414 .332
3	2032 1.12	8649 3.51	1351 56	4434 5.39	1000 5.11	1193 2.01	1236 0.66	1494 6.29	9171 .712	- 1392 2.2	9115 .774	1338 .848
4	3622 4.96	- 3040 .15	4477 5.77	9286 3.64	1491 0.64	1437 5.27	- 1413 .99	6152 .499	- 1392 2.4	2523 .901	1790 3.45	- 1413 .96
5	1790 3.46	4347 .871	1542 6.65	1500 3.55	1730 3.62	1225 4.06	7601 .11	- 1413 .83	- 1392 2.4	3214 3.23	6656 7.46	- 1186 8.5
6	1422 96.6	5972 .48	1542 5.56	1544 1.7	4725 0.63	4724 9.47	- 2409 .3	1293 1.69	- 1392 2.4	1786 .166	1909 60.5	- 1401 .39
7	1588 0.56	- 1032 49	1283 7.99	1657 7.87	8010 .65	2647 0.77	9958 8.25	3329 7.91	1396 4.65	4437 .116	- 5138 1.9	619. 4599
8	- 4608 .86	1994 5.54	2144 4.44	- 7125 7	1515 .405	9410 .836	3305 1.24	6255 7.26	1288 06.9	8569 .392	4290 .078	1099 .996
9	3803 7.67	2262 9.9	1024 0.48	2552 .107	2315 .8	2880 3.15	1540 1.33	2176 2.78	1709 87	2023 0.87	5199 2.77	6405 .727
10	1740 9.04	3469 6.15	6314 .484	6313 .976	1082 .034	1617 .727	1402 88	5068 .181	1981 6.52	1351 56.9	3532 9.84	6314 .491
11	6656 7.96	2428 3.43	6314 .484	6314 .396	752. 6497	1519 .727	4014 9	1020 3.37	4093 8.55	3474 2.2	6656 7.99	4237 .094
12	2005 .757	2841 .323	- 1569 .72	269. 097	2653 .735	- 1082 69	- 3647 6.8	1059 .518	4234 .838	5703 9.18	4214 .089	6638 .813

سپس مقدار R^2 (ضریب تعیین^۱) برای این دو ماتریس محاسبه شد. مقادیر R^2 نشانگر آن است که این تخمین به چه میزان دقت داشته است و هرچه R^2 به عدد ۱ نزدیکتر باشد نشان‌دهنده دقت بالاتر در تخمین و نزدیکتر بودن مقادیر تخمین زده شده با مقادیر اولیه می‌باشد. مقدار R^2 میان ماتریس سفر اولیه و تخمین زده شده با استفاده از داده‌های شهر مشهد برابر با ۰,۵۶ بوده است.

۵-۳-۲ ماتریس سفر تخمین زده شده شهر اصفهان

در این بخش ماتریس سفر اولیه شهر اصفهان مربوط به سال ۱۳۹۵ ارائه و سپس ماتریس سفر تخمین زده شده ارائه خواهند شد. ابتدا در جدول ۵-۴ ماتریس سفر سال ۱۳۹۵ شهر مشهد ارائه شده است و T مففی از کلمه Target به معنای هدف در فرآیند تخمین با شبکه عصبی است.

جدول ۵-۳ ماتریس سفر Target شهر اصفهان

T	1	2	3	4	5	6	7	8	9	10
1	80043	26457	34120	9205	31884	15579	24955	60083	27972	12286
2	27649	108740	10298	2526	10058	5992	7316	54154	13972	2745
3	33309	9994	128391	57810	32509	36283	66456	46817	7715	65048
4	9412	2570	59776	106888	18269	19752	11458	10189	1504	22946
5	33575	9534	35284	18490	447079	99441	21632	33799	9432	13521
6	16211	6071	37019	18917	99471	140845	18475	20739	4351	14511
7	26024	7390	69969	11695	23401	18468	237954	66206	2312	58204
8	62734	54545	50440	10889	35321	21479	65778	409513	8519	19368
9	28553	14140	7569	1564	9150	4099	2122	8330	46746	1694
10	12405	2659	68755	22983	13118	15124	55824	18912	1641	151574

¹ Coefficient of Determination

سپس در جدول ۵-۵ ماتریس سفر تخمین زده شده مشاهده می‌گردد. در این ماتریس O مخفف

کلمه OutPuts به معنای خروجی شبکه عصبی می‌باشد.

جدول ۴-۵ ماتریس سفر OutPut شهر اصفهان

O	1	2	3	4	5	6	7	8	9	10
1	27505.4	37099.48	56214.28	-12132.5	157853.5	14395.79	29562.13	9789.912	6659.7	29344.04
2	18488.75	37390.79	26610.09	-2688.08	125553.6	21581.65	38538.19	45263.76	6415.703	21271.57
3	43611.46	42644.28	82084.07	-7013.52	171426.4	22007.32	44743.49	33212.36	6293.383	38169.68
4	32642.54	30255.21	48705.8	7822.612	136903.6	30659.86	32944.87	27545.73	14928.15	28851.87
5	4234.749	17405.21	36053.37	29639.27	338637.2	128167.3	167856.5	93694.77	2387.378	21921.21
6	30716.42	15393.57	40742.01	13310.08	196287.7	62210.7	26304.35	9286.34	14512.78	29503.37
7	46976.25	23103.26	70105.5	22396.23	190492.2	24673.22	163565.3	189223.6	14262.52	37899.2
8	75841.81	26800.7	74761.4	25242.98	181342.8	28821.65	139525	185836.9	39722.77	12857.46
9	13888.2	28225.17	29384.55	-7111.84	158532.7	18689.66	31250.5	26768.46	8666.688	12495.53
10	42772.95	25907.89	53490.07	17699.57	88220.14	47298.27	29066.49	43730.81	20231.01	53636.07

سپس مقدار R^2 برای این دو ماتریس محاسبه شد و مقدار آن برابر با $0/3$ بوده است که مقادیر

نامناسبی بنظر می‌رسند.

اگرچه مقادیر R^2 در این تحقیق مقادیر بالایی نیست اما باتوجه به آنکه حین فرآیند جمع آوری

داده‌های شهر مشهد، داده‌های کافی در اختیار این تحقیق قرار نگرفت و تنها داده‌های مربوط به

سال ۱۳۹۵ و همان داده‌ها نیز نیاز به تبدیل نواحی ترافیکی به نواحی شهرداری داشتند که از

دقت داده‌ها کاسته خواهد شد و همچنین فایل مربوط به ماتریس سفر و داده‌های سرشماری

جمعیتی پس از آماده سازی بر اساس نواحی سیزده‌گانه و فایل کاربری اراضی بر اساس مناطق

دوازده‌گانه بود لذا در این تحقیق از تعداد سفرهای منطقه سیزده(منطقه ثامن) در فرآیند تخمین

استفاده نشد. و همچنین هنگامی جمع‌آوری داده‌های شهر اصفهان تنها داده‌های مربوط به سال

۱۳۹۵ در اختیار این تحقیق قرار گرفت و در همان داده‌ها نیز در ماتریس سفر داده‌های مربوط به منطقه ۱۵ شهرداری موجود نبود و بالاجبار داده‌های کاربری این ناحیه نیز در تحقیق لحاظ نگردید و تعدادی از نواحی نیز بایکدیگر تلاقی داشتند و در این تحقیق بایکدیگر ادغام شدند و فرض شد آن زوج ناحیه‌ها هر زوج، یک ناحیه است. تمامی این فرضیات باعث کاهش دقت تخمین شد اما همچنان باتوجه به مقادیر بدست آمده پیش‌بینی می‌شود در صورت استفاده از تعداد داده‌های شهرهای دیگر در کنار داده‌های موجود و همچنین داده‌های سال‌های مختلف موجود برای هر یک از شهرها بتوان از تعداد داده‌های بسیار بیشتر و دقیق‌تری جهت تعلیم شبکه عصبی استفاده نمود و بالتبع نتایج بسیار دقیق‌تری حاصل گردد. و با توجه به شرایط داده‌های موجود در این تحقیق، نتایج بدست آمده قابل قبول می‌باشند.

فصل ٦ : نتایج

۶-۱ مقدمه

در این تحقیق هدف آن بود تا بتوان با استفاده از هوش مصنوعی، بر اساس پارامترهای کاربری اراضی و جمعیت حاضر در هر یک از نواحی شهری، تخمین مناسبی از ماتریس سفر آن شهر بدست آید.

ابتدا داده‌های مربوط به سطح زیربنای کاربری اراضی مختلف از سازمان فناوری اطلاعات و ارتباطات شهرداری و ماتریس سفر از سازمان ترافیک دو کلانشهر مذهبی و توریستی مشهد و اصفهان تهیه شدند.

داده‌ها طبقه‌بندی و آماده‌سازی شدند و سپس با بررسی مطالعات پیشین ۴ روش که احتمال تخمین مناسب تری داشتند جهت تخمین ماتریس سفر بر اساس داده‌های موجود، پیشنهاد شدند:

۱- شبکه عصبی جعبه ابزار متلب

۲- شبکه عصبی چندلایه بدون داده‌های اعتبار سنجی

۳- شبکه عصبی با حذف داده‌های نویزی

۴- شبکه عصبی پیشخور

و در هر روش، دو حالت استفاده از داده‌های نرمال‌سازی شده و نرمال‌سازی نشده و در هر حالت پنج دسته‌بندی مختلف بر اساس ضرایب همبستگی ۰/۵ و ۰/۶ و ۰/۷ و ۰/۸ و ۰/۹ بررسی شدند.

۲-۶ نتایج

- داده های مربوط به شهر اصفهان در ۴۰ حالت بررسی شده در ۲۵ مورد دارای MSE کمتر نسبت به شهر مشهد بوده اند و بنظر می رسد علت این باشد که تعداد سفرها در ماتریس سفر شهر مشهد وابستگی بالاتری به آمار مسافریین داشته است درحالی که ماتریس سفر شهر اصفهان بیشتر تحت تاثیر شهروندان و در نتیجه متناسب تر با کاربری اراضی و جمعیت ساکن در نواحی شهری بوده است.
- حذف داده های نویزی تاثیری بر نتایج تخمین ندارد.
- بهترین نتایج MSE در هنگام استفاده از داده های نرمال سازی شده با استفاده از شبکه عصبی چندلایه بدون لحاظ کردن محدودیت اعتبارسنجی داده ها حاصل شد.
- روش استفاده از شبکه عصبی بدون محدودیت داده های اعتبار سنجی از باقی روش های مورد بررسی در این تحقیق بر روی داده های مورد استفاده در این تحقیق نتایج مناسب تری حاصل کرد.
- استفاده از داده های دقیق کاربری اراضی واقع در نواحی شهری و مشخصات جمعیتی آن نواحی می تواند در تخمین ماتریس سفر به کار رود که در صورت استفاده از شبکه عصبی مناسب به نتایج قابل قبولی منجر خواهد شد.
- R^2 ماتریس تخمین زده شده شهر مشهد در مقایسه با R^2 ماتریس تخمین زده شده شهر اصفهان مقدار بالاتری داشت و بنظر می رسد علت آن بوده است که داده های مربوط به شهر مشهد از دقت بالاتری برخوردار بودند درحالی که در آماده سازی داده های شهر اصفهان بعلت موجود نبودن تعداد سفرهای یکی از نواحی شهرداری در محاسبات از آن ناحیه استفاده نشد و همچنین بعلت تعداد بالاتر نواحی (در این تحقیق شهر مشهد ۱۲ و

شهر اصفهان ۱۰ ناحیه در نظر گرفته شده بودند) و در نتیجه تعداد بالاتری داده‌های اولیه جهت آموزش شبکه عصبی داشت. در نتیجه می‌توان بیان داشت در صورت استفاده از داده‌های اولیه دقیق‌تر و همچنین تعداد داده‌های بیشتر، ماتریس سفر بصورت بسیار دقیق‌تری تخمین زده خواهد شد.

۳-۶ پیشنهادات برای ادامه تحقیق:

- تحقیق بر روی داده‌های شهرستان‌ها و کلانشهرهای دیگر انجام شود.
- تعداد هتل‌ها و مسافرخانه‌ها و اماکن شاخص گردشگری بعنوان کاربری جداگانه ای جهت در نظر گرفتن تاثیر مسافرت بر سفرهای شهری در تحقیقات آینده لحاظ گردد.
- برای تهیه داده‌های کاربری در مناطق مختلف شهر از تصاویر ماهواره‌ای استفاده گردد.

مراجع

- [1] Spiess, H. A(1987) maximum-likelihood model for estimating origin-destination matrices. IEEE Trans. Pattern Anal. Mach. Intell., 24(9):1167–1183.
- [2] Nguyen, Sang and Cascetta, Ennio. (1987) A unified framework for estimation or updating origin destination matrices from traffic counts. Tech. Rep. PUB-580, Université de Montréal (CA).
- [3] Hazelton, Martin L. (2000) Estimation of origin-destination matrices from link flows on uncongested networks. Transportation Research Part B: Methodological, 34(7):549–566.
- [4] Fisk, C. S. (1988) On combining maximum entropy trip matrix estimation with user optimal assignment. Transportation Research Part B: Methodological, 22(1):69–73.
- [5] Wilson, A.G. (2011) Entropy in Urban and Regional Modelling. Monographs in spatial and environmental systems analysis. Routledge.
- [6] Xie, Chi, Kockelman, Kara M., and Waller, S. Travis. (2011) A maximum entropy least squares estimator for elastic origin destination trip matrix estimation. Transportation Research Part B: Methodological, 45(9):1465–1482.
- [7] Willumsen, L.G., of Leeds. Institute for Transport Studies, University, and Britain, Science Research Council Great. (1978) Estimation of an O-D Matrix from Traffic Counts: A Review. Working paper: Institute of Transport Studies. Institute for Transport Studies, University of Leeds.
- [8] Brenninger-Göthe, Maud, Jörnsten, Kurt, and Lundgren, Jan T. (1989) Estimation of origin-destination matrices from traffic counts using multiobjective programming formulations. Transportation Research Part B: Methodological, 23(4):257–269.
- [9] Malouf, Robert. (2002) A comparison of algorithms for maximum entropy parameter estimation..
- [10] Van Zuylen, Henk J. and Willumsen, Luis G. (1980) The most likely trip matrix estimated from traffic counts. Transportation Research Part B: Methodological, 14(3) : 281–293.
- [11] Gu, Yu, McCallum, Andrew, and Towsley, Don. (2005) Detecting anomalies in network traffic using maximum entropy estimation. pp. 32–32.
- [12] Bell, Michael G. H. (1991) The estimation of origin-destination matrices by constrained generalised least squares. Transportation Research Part B: Methodological, 25(1):13–22.
- [13] Bierlaire, M. and Toint, Ph.L. Meuse: (1994) an origin-destination matrix estimator that exploits structure.
- [14] Cascetta, Ennio. (1984) Estimation of trip matrices from traffic counts and survey data: A generalized least squares estimator. Transportation Research Part B: Methodological , 18(4-5):289–299.

-
- [15] Nie, Yu, Zhang, H.M., and Recker, W.W. (2005) Inferring origin-destination trip matrices with a decoupled gls path flow estimator. *Transportation Research Part B: Methodological*, 39(6):497–518.
- [16] Yang, Hai, Iida, Yasunori, and Sasaki, Tsuna. (1994) The equilibrium-based origin-destination matrix estimation problem. *Transportation Research Part B: Methodological*, 28(1):23–33.
- [17] Judge, George G., Griffiths, William E., Hill, R. Carter, and Lee, Tsoung-Chao, eds. (1980) *The theory and practice of econometrics*. Wiley series in probability and mathematical statistics. Wiley, New York [u.a.].
- [18] Snickars, F. and Weibull, J. W. (1977), “A minimum information principle, theory and practice”, *Regional Science and Urban Economics*, 7.
- [19] Wilson, A. G. (1970) *Entropy in Urban and Regional Modeling*, Methuen, Inc., New York.
- [20] Predicting travel mode of individuals by machine learning, Hichem Omrani, *Transportation Research Procedia*, Volume 10, Pages 840-849, 2015
- [21] Forecasting origin-destination matrices by using neural network approach: A comparison of testing performance between back propagation, variable learning rate and levenberg-marquardt algorithms , Gusri Yaldi, Michael A P Taylor, Wen Long Yue , 2011
- [22] SHORT TERM URBAN TRAFFIC FORECASTING USING DEEP LEARNING, G. Albertengo , W. Hassan, 2018.

Abstract

One of the most important parameters in transportation studies and management is the number of trips between origin and destination, which is shown using the travel matrix. Because preparing a trip matrix requires a lot of time and money, in recent years, several methods have been developed and used to estimate the travel matrix. The aim of this study is to use artificial intelligence to provide a suitable estimate of the travel matrix using the population living in the areas of origin and destination and the total area of land use infrastructure located in each area. For this purpose, previous researches were reviewed and it was concluded that a research similar to the subject of this research can be completed using artificial intelligence. After that, the data related to the population census as well as the total area of land use infrastructure in the municipal areas in the cities of Mashhad and Isfahan were collected. Using methods 1) MATLAB 2013 software toolbox neural network 2) Multilayer neural network without limitation of validation data 3) Neural network with noise data removal 4) Feed-forward neural network were used to estimate the travel matrix.

In this study, the most appropriate results were obtained when using normalized data in a multilayer neural network without limitation of validation data.

Keywords: Travel, Trip matrix, Estimation, Artificial intelligence, Neural network, MATLAB 2013, Data normalization



Faculty of Civil Engineering

M.Sc. Thesis in Road and Transport Engineering

**estimation of origin-destination matrix by using
artificial intelligence**

By: Abolfazl Sajadi

Supervisor:

Dr. Hosein Ghasemzadeh Tehrani

October 2021